MODELING OF TRANSIENT PROTEIN-PROTEIN INTERACTIONS: A STRUCTURAL STUDY OF THE THIOREDOXIN SYSTEM

A Thesis Submitted to the College of Graduate Studies and Research in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in the Department of Chemistry
University of Saskatchewan
Saskatoon, Canada

Josiah Obiero
Fall, 2010

Department of Chemistry
University of Saskatchewan

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work, or in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of materials in this thesis in whole or in part should be addressed to:

Head of the Department of Chemistry

University of Saskatchewan

110 Science Place

Saskatoon, Saskatchewan

Canada, S7N 5C9

ABSTRACT

Protein-protein interactions play a central role in most biological processes. One such biological process is the maintenance of a reducing environment inside the cell. To maintain an internal reducing environment, living cells have evolved two enzymatic systems (glutathione and thioredoxin (Trx) systems). The Trx system is composed of the enzyme TrxR and its substrate Trx. The two proteins constitute an important thiol-dependent redox system that catalyzes the reduction of many proteins that are responsible for a variety of cellular functions. The system relies on transient protein-protein interactions between Trx and TrxR for its function.

Cross-reactivity of components of the Trx system between species has been shown to be medically relevant. For example, *Helicobacter pylori* Trx (HP Trx) is thought to mediate catalytic reduction of human immunoglobulins and thus facilitate immune evasion. It has also been proposed that *Helicobacter pylori* gains access to the impenetrable gastric mucous layer by using secreted HP Trx to reduce the disulfide bonds present in the cysteine-rich mucin regions that are responsible for cross-linking mucin monomers. Therefore, disruption of secreted HP Trx-host protein interaction may result in restoration of the viscoelastic and hydrophobic protective properties of mucus. Previous studies aimed at understanding the nature of cross-reactivity of Trx system components among various species have shown that Trxs have higher affinity for cognate TrxRs (same species), than for TrxRs from different species. However, the basis for this specificity is not known. A growing body of evidence suggests that most protein-protein interactions are mediated by a small number of protein-protein interface residues, referred to as “hot spot” residues or binding epitopes. Therefore, understanding the biochemical

basis of the affinity of proteins for their partners usually begins by identifying the hot spot residues responsible for the protein complex interactions.

In this study, the crystal structures of *Deinococcus radiodurans* thioredoxin reductase (DR TrxR) and *Helicobacter pylori* TrxR (HP TrxR) were determined at 1.9 Å and 2.4 Å respectively. Analysis of the Trx-binding sites of both structures suggests that the basis of affinity and specificity of Trx for TrxR is primarily due to the shape rather than the charge of the surface. In addition, the complex between *Escherichia coli* thioredoxin reductase (EC TrxR) and its substrate thioredoxin (EC Trx) was used to identify residues that are responsible for TrxR-Trx interface stability. Using computational alanine scanning mutagenesis and visual inspection of the EC TrxR-Trx interface, 22 EC TrxR side chains were shown to make contact across the TrxR-Trx interface. Although more than 20 EC TrxR side chains make contact across the TrxR-Trx interface, our results suggest that only 4 residues (F81, R130, F141, and F142) account for the majority of the EC TrxR-Trx interface stability. Individual replacement of equivalent DR TrxR residues (M84, K137, F148, F149) with alanine resulted in drastic changes in binding affinity, confirming that the four residues account for most of TrxR-Trx interface stability. These hot spot residues are surrounded by less important residues (hydrophobic and hydrophilic) that are also predicted to contribute to interface stability. F148 and F149 are invariant across bacterial TrxRs, however other residues that contact Trx are less conserved including M84 and K137. When M84 and K137 were changed to match equivalent *E. coli* TrxR residues (K137R, M84F); *D. radiodurans* TrxR substrate specificity was altered from its own Trx to that of *E. coli* Trx. The results suggest that a small subset of the TrxR-Trx interface residues are responsible for the majority of Trx binding affinity and specificity, a property that has been shown to general to protein-protein interfaces.

ACKNOWLEDGEMENTS

It is my pleasure to express my sincere thanks to my supervisor Dr. David A. R. Sanders for accepting me as a graduate student in his lab, and for his indispensable guidance and patience during my tenure as a graduate student. He has given me unequivocal support over the years especially when it seemed like nothing was working. I also owe my sincerest thanks to our collaborator and my advisory committee member, Dr. David Palmer, for his constant assistance during my graduate program, insightful discussions, and for allowing me to use facilities in his lab. In addition, I would like to thank my other supervisory committee members, Dr. Yu Luo and Dr. Matthew Paige, for their encouragement and support. Talking to them is always an enjoyable thing.

I wish to offer my warm thanks to the past and present members of our lab and Dr. Palmer's lab: Karen, Hongyang, Fang, Vennessa, Meghan, Sarah, Chris, Dr. van Stratten, and Dr. Sheoran. They made our lab an enjoyable work place. Special thanks to Eshwari, Salina and KP for their support and willingness to share their expertise with me over the years. I am also grateful to the current and past Department of Chemistry staff: Dr Ravindran, Dr. MacKenzie, Dr. Bertole-Scott, Cynthia, Tanis, Virginia, Jason, and Cathy for their support with my teaching and research program.

Most importantly I thank God for life, a family of friends, and my real family. Your presence and support mean more to me than you will ever know. Eliud, 'Igwe' Denis, Catharine, 'McCain' Donkuru, Abdalla, Hamid 'the Arab', Riana, Jackson, Barbara, Lilian, auntie Carol, uncle Da & Ziyanda, uncle Malone & Del, uncle Patrick & Hilda, uncle Max & Pearl, uncle Passmore & Njekwa, Carla, Velma, Cyprian, Allison, Glyne & Brenda, Abednego, Joseph, Sarah, Kegan, Carmen, Shannon, Inonge, Samu, Alex, Michele, and Danielli - you are all the best!

DEDICATION

To my friend Dr. Phil Raney (1954-2007)

And to

My parents, Samson and Rhoda Obiero

Thanks for your DNA!

LIST OF ABBREVIATIONS

AADP	Aminopyridine adenine dinucleotide phosphate
AEBSF	[4-(2- Aminoethyl) benzenesulfonyl fluoride]
A. thaliana	*Arabidopsis thaliana*
ASA	Solvent accessible surface area
APS	Ammonium persulfate
BSA	Bovine serum albumin
CAPRI	Critical assessment of predicted interactions
CHARMM	Chemistry at HARvard Molecular Mechanics
CM-cellulose	Carboxymethyl-cellulose
DEAE-cellulose	Diethylaminoethyl-cellulose
D. radiodurans	*Deinococcus radiodurans*
DR Trx	*Deinococcus radiodurans* Thioredoxin
DR TrxR	*Deinococcus radiodurans* Thioredoxin reductase
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DOC	Deoxycholate
DPDS	4,4'-dithio-pyridine
DSB	Double stranded breaks
DTNB	[5, 5'- Dithio-bis (2-nitrobenzoic acid)]
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EC Trx	*Escherichia coli* Thioredoxin

LIST OF ABBREVIATIONS – *Continued*

EC TrxR	*Escherichia coli* Thioredoxin reductase
EDTA	Ethylenediamine tetra-acetic acid
FAD	Flavin adenine dinucleotide
FO	Flavin oxidizing conformation
FR	Flavin reducing conformation
GV	Gap volume index
Gy	Gray, the SI unit of absorbed radiation dose of ionizing radiation
H. pylori	*Helicobacter pylori*
HP Trx	*Helicobacter pylori* Thioredoxin
HP TrxR	*Helicobacter pylori* Thioredoxin reductase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HMW	High molecular weight
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl- β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kDa	Kilo Dalton
LB	Luria-Bertani
LLG	Log likelihood gain
LMW	Low molecular weight
MW	Molecular weight
MrBUMP	Molecular replacement by use of multiple proteins
NADPH	Nicotinamide adenine dinucleotide phosphate (Reduced form)

LIST OF ABBREVIATIONS – *Continued*

NMR	Nuclear magnetic resonance
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PFU	*Pyrococcus furiosus*
PMA	Phenyl mercuric acetate
R. capsulatus	*Rhodobacter capsulatus*
Rmsd	Root mean square deviation
RNA	Ribonucleic acid
ROS	Reactive oxygen species
Rpm	Revolution per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SSSC	Saskatchewan Structural Sciences Centre
TEMED	N, N, N', N'-Tetramethylethylenediamine
TNB	5-Thio-2-nitrobenzoic acid
Tris	*Tris*(hydroxymethyl)aminomethane
Tpx	Thioredoxin peroxidase
Trx	Thioredoxin

TrxR	Thioredoxin reductase

TABLE OF CONTENTS

PERMISSION TO USE i
ABSTRACT ii
ACKNOWLEDGEMENTS iiv
DEDICATION v
LIST OF ABBREVIATIONS vi
TABLE OF CONTENTS x
LIST OF TABLES xiv
LIST OF FIGURES xv
CHAPTER ONE: INTRODUCTION 1
1.1 PROTEIN-PROTEIN INTERACTIONS 1
1.2 THIOREDOXIN SYSTEM 5
1.2.1 General Background 5
1.2.2 Structure of Trx 9
1.2.3 Structure of TrxR 10
1.2.4 Conformational States of TrxR 11
1.3 CHARACTERIZATION OF PROTEIN-PROTEIN INTERACTIONS…….16
1.3.1 Enzyme Kinetics 16
1.3.2 Protein-Protein Docking 19
1.4. OVERVIEW OF PROTEIN CRYSTALLOGRAPHY 21
1.4.1 Protein Crystal Growth 21
1.4.2 X-rays and Diffraction Theory 23
1.4.3 Symmetry 25
1.4.4 The Structure Factor Equation 26
1.4.5 The Phase Problem 27
1.4.6 Structure Solution by Molecular Replacement 28
1.4.7 Model Building, Refinement, and Validation 29
1.5 MODEL SYSTEMS 31
1.5.1 *D. radiodurans* 31
1.5.2 *H. pylori* 33
1.6 INDUSTRIAL AND PHARMACEUTICAL APPLICATIONS OF THE THIOREDOXIN SYSTEM 34
1.7 AIMS AND OBJECTIVES 35

TABLE OF CONTENTS - Continued

CHAPTER TWO: MATERIALS AND METHODS 37
2.1 PROTEIN OVEREXPRESSION 37
2.2 PROTEIN PURIFICATION 38
2.2.1 Purification of HP Trx1 38
2.2.2 Purification of HP Trx2 39
2.2.3Purification of HP TrxR 39
2.2.4 Purification of DR TrxR and Mutants 40
2.2.5 Purification of DR Trx1 41
2.2.6 Purification of EC TrxR 42
2.2.7 Purification of EC Trx 42
2.3 DETERMINATION OF PROTEIN PURITY 42
2.4 ESTIMATION OF PROTEIN CONCENTRATION 43
2.5 PREPARATION OF MUTANTS 44
2.6 ACTIVITY ASSAYS 47
2.6.1 Insulin Reduction Assay 47
2.6.2 DTNB Reduction Assay 48
2.7 FLUORESCENCE MEASUREMENTS 48
2.8 THIOREDOXIN REDUCTASE-THIOREDOXIN COMPLEX FORMATION……49
2.9 CRYSTALLIZATION AND CRYSTAL HARVESTING 50
2.9.1 HP TrxR 50
2.9.2 DR TrxR 50
2.10 DATA COLLECTION AND PROCESSING 50
2.11 HOMOLOGY MODELING 51
2.12 PROTEIN-PROTEIN DOCKING 52
2.13 CALCULATION OF INTERACTION ENERGIES 52
CHAPTER THREE: RESULTS AND DISCUSSION 54
3.1 PROTEIN OVER-EXPRESSION, PURIFICATION AND CHARACTERIZATION 54
3.1.1 Protein Over-Expression and Purification 54
3.1.2 Estimation of Protein Concentration 55

TABLE OF CONTENTS - Continued

3.2 ENZYME KINETICS 55
3.2.1 Activity Dependence on Cofactor Concentration 55
3.2.2 Activity of DR Trx1 57
3.2.3 Steady-state Kinetic Analysis of DR TrxR and DR Trx1 59
3.2.4 Cofactor Specificity 62
3.3 CRYSTALLIZATION 64
3.3.1 Crystallization of DR TrxR 64
3.3.2 Crystallization of HP TrxR 66
3.4 DATA COLLECTION AND PROCESSING 67
3.5 STRUCTURE DETERMINATION 69
3.5.1 DR TrxR Structure Determination 70
3.5.2 HP TrxR Structure Determination 71
3.6 MODEL BUILDING, REFINEMENT AND VALIDATION 71
3.7 THIOREDOXIN REDUCTASE STRUCTURES 72
3.7.1 Oxidized DR TrxR Structure 72
3.7.2 Oxidized HP TrxR Structure 76
3.7.3 FAD Binding 77
3.7.4 NADPH binding 81
3.7.5 Dimer Interface 83
3.7.6 Comparison with other Bacterial TrxR Structures 86
3.8 CONFORMATIONAL STATES OF THIOREDOXIN REDUCTASE 88
3.8.1 Fluorescence Spectroscopy 88
3.8.2 Complex Formation 90
3.9 THIOREDOXIN REDUCTASE SPECIES SPECIFICITY 93
3.9.1 Cross Reactivity of the DR Trx System with the EC Trx System 93
3.9.2 Structural Analysis of the Trx Binding Site 95
3.9.3 Modeling of TrxR-Trx Complex Structures 96
3.9.4 Identification of Hot Spot Residues 102
3.9.5 Kinetic Analysis of Hot Spot Residues 105
3.9.6 Comparison of Calculated TrxR-Trx Interaction Energies versus Affinities 108
3.9.7 Design of DR TrxR for Altered Trx Specificity 110

TABLE OF CONTENTS - Continued

CHAPTER FOUR: CONCLUSIONS AND PERSPECTIVES 112
4.1 CONCLUSIONS 112
4.2 FUTURE DIRECTIONS 114
APPENDIX 117
A) CHARACTERIZATION OF THIOREDOXIN-THIOREDOXIN REDUCTASE INTERACTIONS BY ISOTHERMAL TITRATION 117
A1: Introduction 117
A2: Materials and Methods 120
A3: Results and Discussion 120
B) CRYSTALLIZATION OF HP TRX1 AND HP TRX2 123
B1: Introduction 123
B2: Materials and Methods 124
B2.1: Crystallization of HP Trx1 124
B2.2: Crystallization of HP Trx2 124
B3: Results and Discussion 125
REFERENCES 127

LIST OF TABLES

Table 2.1: Overexpression and induction conditions *H. pylori*, *E. coli*, and *D. radiodurans* proteins 38

Table 2.2: BSA standards samples for Bradford assay 43

Table 2.3: List of primers used for site-derected mutagenesis 46

Table 2.4: PCR conditions for mutagenesis 47

Table 3.1: Kinetic constants for DR Trx system 61

Table 3.2: Data collection and refinement statistics for DR TrxR and HP TrxR 68

Table 3.3: K_m values of homologous and heterologous Trx sytems 93

Table 3.4: Surface area and gap volume indices 97

Table 3.5: Computational alanine scanning mutagenesis results from Roberta server (TrxR) .. 104

Table 3.6: Kinetic parameters of Trx binding site mutants of TrxR 106

Table 3.7: Kinetic parameters of DR TrxR double mutant (K137R, M84F) 111

Table 4.1: Computational alanine scanning mutagenesis results from roberta server (Trx) 116

Table B.1: Kinetic constants for *H. pylori* Trx system 123

Table B.2: Data and statistics for Trx2 126

LIST OF FIGURES

Figure 1.1: Movement of reducing equivalents from NADPH to the active site disulfide of TrxR via $FADH_2$ 7

Figure 1.2: Schematic representation of TrxR conformation during catalysis 8

Figure 1.3: Structure of EC Trx 9

Figure 1.4: Structure of DR TrxR 10

Figure 1.5: Schematic representation of the FO and FR conformations of TrxR 12

Figure 1.6: Structures of FR and FO conformations of TrxR 14

Figure 1.7: Position of the active site disulfide relative to the isoalloxazine ring of FAD 15

Figure 1.8: Michaelis-Menten plot 19

Figure 1.9: Solubility phase diagram for crystallization from a solution 22

Figure 1.10: Constructive interference conditions 24

Figure 2.1: Confirmation of mutagenesis by DNA sequence alignment 45

Figure 3.1: SDS-PAGE of purified samples of various TrxR and Trx proteins 54

Figure 3.2: Plot of initial velocities vs cofactor concentration 56

Figure 3.3: Plot of DTT dependent reduction of insulin 57

Figure 3.4: Sequence alignment of various Trxs 58

Figure 3.5: Electron flow from NADPH to insulin or DTNB 59

Figure 3.6: Michelis-Menten plot of DR TrxR using DR Trx1 as a substrate 61

Figure 3.7: Sequence alignment of DR TrxR and EC TrxR 63

Figure 3.8: Insulin reduction sssay using NADPH and NADH as a cofactor 64

Figure 3.9: Picture of DR TrxR crystal 65

Figure 3.10: Picture of HP TrxR crystal 66

LIST OF FIGURES – *Continued*

Figure 3.11: Diffraction pattern of DR TrxR 67

Figure 3.12: Electron density map of the initial solution generated by MrBUMP 70

Figure 3.13: Electron density map of the active site disulfide 73

Figure 3.14: The Ramachandran plot of DR TrxR 74

Figure 3.15: NADP(H)-binding domain of DR TrxR 75

Figure 3.16: The Ramachandran plot of HP TrxrR 77

Figure 3.17: Stereoview of key residues that interact with FAD in the DR TrxR structure. 79

Figure 3.18: Sequence alignment of various TrxRs showing thioredoxin binding, nuclotide binding, and dimer formation residues 80

Figure 3.19: HP TrxR structures showing NADP(H) diphosphate binding residues 82

Figure 3.20: Stereo ribbon reprentation of the dimer of DR TrxR 84

Figure 3.21: HP TrxR and DR TrxR dimer interface 85

Figure 3.22: Superimposed TrxR monomers (FO conformation) 89

Figure 3.23: Emission spectra of wild type DR TrxR vs C145S mutant 87

Figure 3.24: Schematic representation of TrxR –Trx complex conformation 91

Figure 3.25: SDS-PAGE of DR TrxR-HP Trx2 complex 92

Figure 3.26: Superposition of various TrxR structures 96

Figure 3.27: Electrostatic surface potentials of various TrxRs and Trxs 101

Figure 3.28: Plot of experimetal K_m vs E 109

Figure 4.1: Computational alanine scanning mutagenesis results from roberta server (TrxR) .. 113

Figure 4.2: Amino-acid sequence alignments of Trxs from various prokaryotic and eukaryotic sources 115

Figure A.1: Basic design of an ITC instrument 118

Figure A.2: An ITC thermogram of titration of RNase with 3’ CMP 119

Figure A.3: An ITC thermogram of titration of EC TrxR with EC Trx1 122

Figure B.1: Trx2 crystals 125

CHAPTER ONE: INTRODUCTION

1.1 PROTEIN-PROTEIN INTERACTIONS

Protein-protein interactions play a key role in all cellular processes; therefore, understanding how proteins interact with each other is the subject of intense research (Moreira et al, 2007). Protein-protein interactions can be broadly classified permanent or transient (Devos & Russell, 2007). Permanent protein-protein complexes facilitate transport (e.g. the nuclear pore), transcription (e.g. RNA polymerases), translation (e.g. the ribosome), metabolism (e.g. glutamine synthetase), signaling (e.g. G-protein complexes in eukaryotes) (Devos & Russell, 2007). Transient protein-protein interactions are also important in many biological processes including signal transduction (e.g. hormone-receptor), gene transcription (e.g. Thioredoxin (Trx)-transcription factors), DNA synthesis (e.g. Trx-ribonucleotide reductase) and regulation of the redox status of the cell (e.g. Thioredoxin-thioredoxin peroxidase (Trx-Tpx)) (Clackson & Wells, 1995; Lo Conte et al, 1999; Mustacich & Powis, 2000). Non-physiological protein-protein interactions such as crystal contacts (protein crystallization) can also be considered as transient protein-protein interactions because, such interactions are non-specific low affinity interactions (Carugo & Argos, 1997).

Typical protein- protein interfaces are usually about 1600 $Å^2$ (~800 $Å^2$ per monomer). Contacts at protein-protein interfaces can be of any type: covalent (disulfides), hydrophobic, electrostatic, or hydrogen bonding (Lo Conte et al, 1999). These contacts can be mediated by side-chain or main-chain interactions (Jackson, 1999). Those interactions that are mediated by side chains tend to have 10 or more side chains that make contact across the interface (Janin & Chothia, 1990; Sheinerman et al, 2000). Unlike folded proteins which are hydrophobic at the core and

hydrophilic at the surface, protein-protein interfaces tend to have both hydrophobic and hydrophilic residues at the interface (Sheinerman et al, 2000). Although the general principles guiding protein-protein interactions are common to permanent and transient protein-protein interfaces, there are some differences between permanent and transient protein-protein interactions. In general, transient protein-protein interfaces tend to have less van der Waals surfaces complimentarity compared to permanent protein-protein interfaces (Jones & Thornton, 1996). In addition, permanent protein-protein interfaces tend to be more hydrophobic than transient protein-protein interfaces, because each partner of transient protein-protein complexes exists independently most of the time (binding surfaces exposed to solvent most of the time) (Jones & Thornton, 1996).

The basic physical principles (thermodynamics) that govern protein-protein interactions are similar to those that govern protein folding. Therefore, the affinity of two uncomplexed proteins that form a noncovalent complex is determined by the change in enthalpy and entropy of the system. Specific interactions (hydrogen bonding, electrostatic and van der Waals forces) and non-specific effects (the hydrophobic effect and conformational entropy) are used to describe the process of complexation (Ajay & Murcko, 1995; Bogan & Thorn, 1998). The protein-protein interaction system consists of the protein molecules and solvent before complex formation, and the complex and solvent after complexation (Ajay & Murcko, 1995). Upon complexation there is a gain in entropy due to reorganization of solvent molecules (expulsion of water molecules from hydrophobic interface residues), a loss in entropy due to association (loss of translational and conformational freedom), and a decrease in enthalpy due to protein-protein interactions (hydrogen bonding, electrostatic and van der Waals forces). Although there is some loss of entropy due to association, the overall process of complexation is spontaneous because of solvent

reorganization and enthalpic changes due to protein-protein interactions (Bogan & Thorn, 1998; Moreira et al, 2007).

Because most cellular functions involve protein-protein interactions, aberrant protein-protein interactions often result in disease. Examples of diseases caused by aberrant protein-protein interactions include cancer (Rudolph, 2007), Parkinson's disease (Suzuki, 2006) and Huntington's disease (Horn et al, 2006). Therefore, disrupting aberrant protein-protein interactions offers a huge therapeutic potential. However, discovering small molecules that can disrupt protein-protein interactions is considered a major challenge for many reasons (Che et al, 2006; Wells & McClendon, 2007). Protein-protein interactions typically occur over a large surface area (about 800 $Å^2$ per monomer) and involve amino acid residues that are not contiguous in the polymer chain. Eight hundred $Å^2$ is a much larger surface area than a small molecule can cover (typically 200 $Å^2$). In addition, the binding surfaces at protein-protein interfaces tend to be relatively flat, often lacking crevices and pockets that provide snug binding sites for small molecules (Che et al, 2006; Wells & McClendon, 2007). Recent success stories, however, suggest that disrupting protein–protein interfaces might be more tractable than had been previously thought. For example, small molecules have been identified that disrupt some protein-protein interactions: human protein double minute 2-tumor suppressor protein p53 (Vassilev et al, 2004), human papilloma virus transcription factor E2-viral helicase E1 (Wang et al, 2004), and tumor necrosis factor-tumor necrosis factor receptor 1 (He et al, 2005).

An important observation from a careful analysis of protein-protein interfaces is that most of the binding energy at protein-protein interfaces is due to a few 'hot spot' residues (functional epitope) rather than all of the interface residues. This phenomenon was first observed by analysis of free energy changes upon mutation to alanine at the interface of human growth hormone and

its receptor (Clackson & Wells, 1995). Hot spot residues have since been observed in several protein-protein complexes (Bogan & Thorn, 1998; Moreira et al, 2007). Although there is no universal definition of hot spot residues, they have been defined as those sites where alanine mutations cause a 1 kcal (Kortemme & Baker, 2002) or 2 kcal (Moreira et al, 2007) increase in free energy (ΔG). Since most of the binding affinity at protein-protein interfaces is due to a small subset of interface residues (binding "hot spots" that are usually clustered at the center of the protein-protein interface), it is not necessary for a small molecule to cover the entire protein-protein interface. A small molecule that maintains the three-dimensional arrangement of the hot spot residues can disrupt protein-protein interactions because it retains the features that confer affinity to protein-protein interfaces. Binding hot spots thus provide a rational basis for design of small molecule inhibitors ("protein-surface mimetics") of protein-protein interactions. As a result, the first step towards protein-surface mimetics is usually to identify hot spot residues responsible for protein complex recognition (Che et al, 2006; Wells & McClendon, 2007).

A systematic analysis of hot spot residues reveals that their distribution is non-random i.e. certain amino acids (Y, W, and R) have a higher propensity than other amino acids for the interface, and that they tend to cluster at the centre of the interface rather on the edges (Moreira et al, 2007). It was further observed that hot spot residues were sheltered from bulk solvent. However, many other interface residues that are occluded from solvent are not hot spot residues, i.e. occlusion of solvent was necessary but not a sufficient determinant of hot spot residues. Indeed a survey of interface residues showed that the amount of change of total side-chain solvent-accessible surface area (ΔASA) on complex formation is a poor predictor of hot spot residues because other factors such as complementarity of van der Waals surfaces contribute to interface stability

(Bogan & Thorn, 1998; Moreira et al, 2007). These observations led to the proposal of the O-ring hypothesis as the basis of the hot spot residue phenomenon (Bogan & Thorn, 1998).

The O-ring hypothesis suggests that hot spot residues are surrounded by energetically less important residues, whose role is to occlude bulk solvent from hot spot residues. Occluding solvent from hot spot residues generates suitable dielectric constant and solvation conditions for a hot spot of interaction energy i.e. a lower effective dielectric constant increases the interaction strength of electrostatic and hydrogen bonding interactions (Bogan & Thorn, 1998). The O-ring phenomenon may explain the observed interface amino acid propensities. Y, W, and R are the most common amino acids at the interface, because they are capable of hydrogen bonding and hydrophobic interactions. The O-ring hypothesis is now widely accepted, although there is no conclusive evidence for it. The assumption that non-hot-spot residues do not participate in important interactions is only valid if it can be demonstrated that there are no secondary effects upon mutation to alanine including conformational changes of side chains, movement of the protein backbone, and reorganization of solvent structure near mutation sites that can compensate for loss of side chain interactions (Moreira et al, 2007; Sanders, 1998).

1.2 THIOREDOXIN SYSTEM

1.2.1 General Background

Proteins in the extracellular environment are rich in stabilizing disulfides, made possible by oxidizing conditions in the extracellular environment. In contrast, most proteins inside the cell contain free sulfhydryl groups, due to the reducing environment inside the cell. The Trx system is one of the two major systems responsible for maintaining a reducing environment inside the cell; the other is the glutathione reductase system (Arner & Holmgren, 2000). All aerobic organisms generate reactive oxygen species (ROS) during normal cell function through

processes such as respiration, or in response to various stimuli such as Ultra-violet (UV) irradiation and chemical carcinogens. However, excessive production of ROS can damage cellular macromolecules like DNA and proteins. Therefore, living cells have evolved various enzymatic systems that scavenge ROS. Thioredoxin-thioredoxin dependent peroxidase (Trx-Tpx) is one such enzymatic system. The Trx system thus plays an essential role in maintaining the integrity of cells (Adler et al, 1999; Nordberg & Arner, 2001). Indeed, gene knock out studies have shown that the Trx system is essential for the viability of some organisms (Krnajski et al, 2002).

The system is composed of Trx and the associated enzyme thioredoxin reductase (TrxR). It is found universally in both prokaryotes and eukaryotes. TrxR catalyzes the transfer of reducing equivalents from NADPH to Trx via a permanently bound FAD and an active site disulfide (Mustacich & Powis, 2000). Electrons are transferred to FAD from NADPH, then from FAD to the active-site disulfide (Figure 1.1) and finally from TrxR to Trx (Figure 1.2). Trx in turn reduces other proteins such as Tpx and ribonucleotide reductase (Lennon & Williams, 1996; Mustacich & Powis, 2000).

There are two distinct forms of this enzyme; low molecular weight thioredoxin reductase (LMW TrxR) (35 kDa) found in prokaryotes and some eukaryotes, and high molecular weight thioredoxin reductase (HMW TrxR) (55 kDa) found in eukaryotes. The two types of TrxR have some differences in structure and mechanism. The structure and mechanism of HMW TrxR has been described in several publications (Zhon et al, 1998; Zhong et al, 2000). Only LMW TrxRs were of interest to this project. In HMW TrxR, NADPH binds in close proximity to the isoalloxazine ring of FAD, allowing electrons to flow from NADPH to active-site disulfide via the isoalloxazine ring of FAD without a major conformational change in the enzyme (Zhong et

al, 2000). In contrast, the LMW TrxR's FAD domain is too far in space from the NADPH domain for electron transfer to occur. It was postulated and later confirmed by the *Escherichia coli* thioredoxin reductase-thioredoxin (EC TrxR- EC Trx) complex structure that during catalysis the NADPH domain rotates 67º while the FAD domain remains fixed, bringing the NADPH domain close to the FAD domain in order for electron transfer to take place (Lennon et al, 2000). Electron transfer then occurs from NADPH to FAD and finally to the active site disulfide. The enzymatic reaction scheme for the LMW Trx system (Lennon & Williams, 1996) is shown in Figure 1.1 and Figure 1.2.

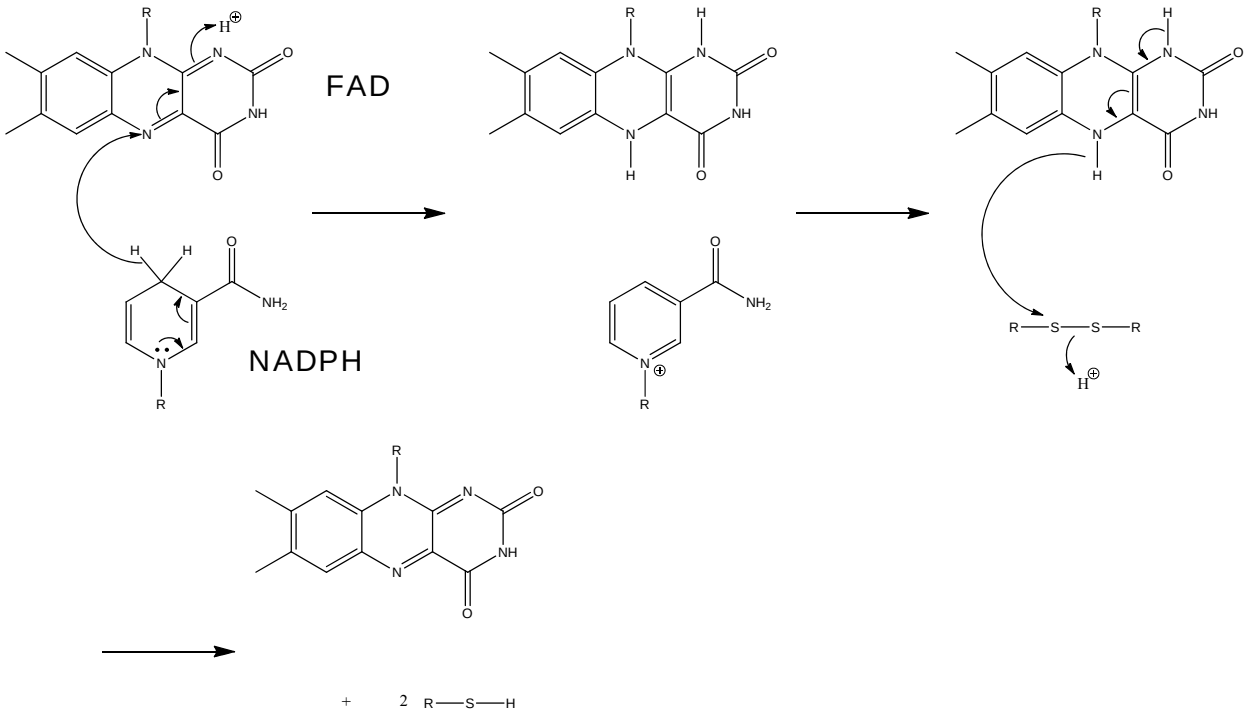


Figure 1.1: Movement of reducing equivalents from NADPH to the active site disulfide of TrxR via $FADH_2$

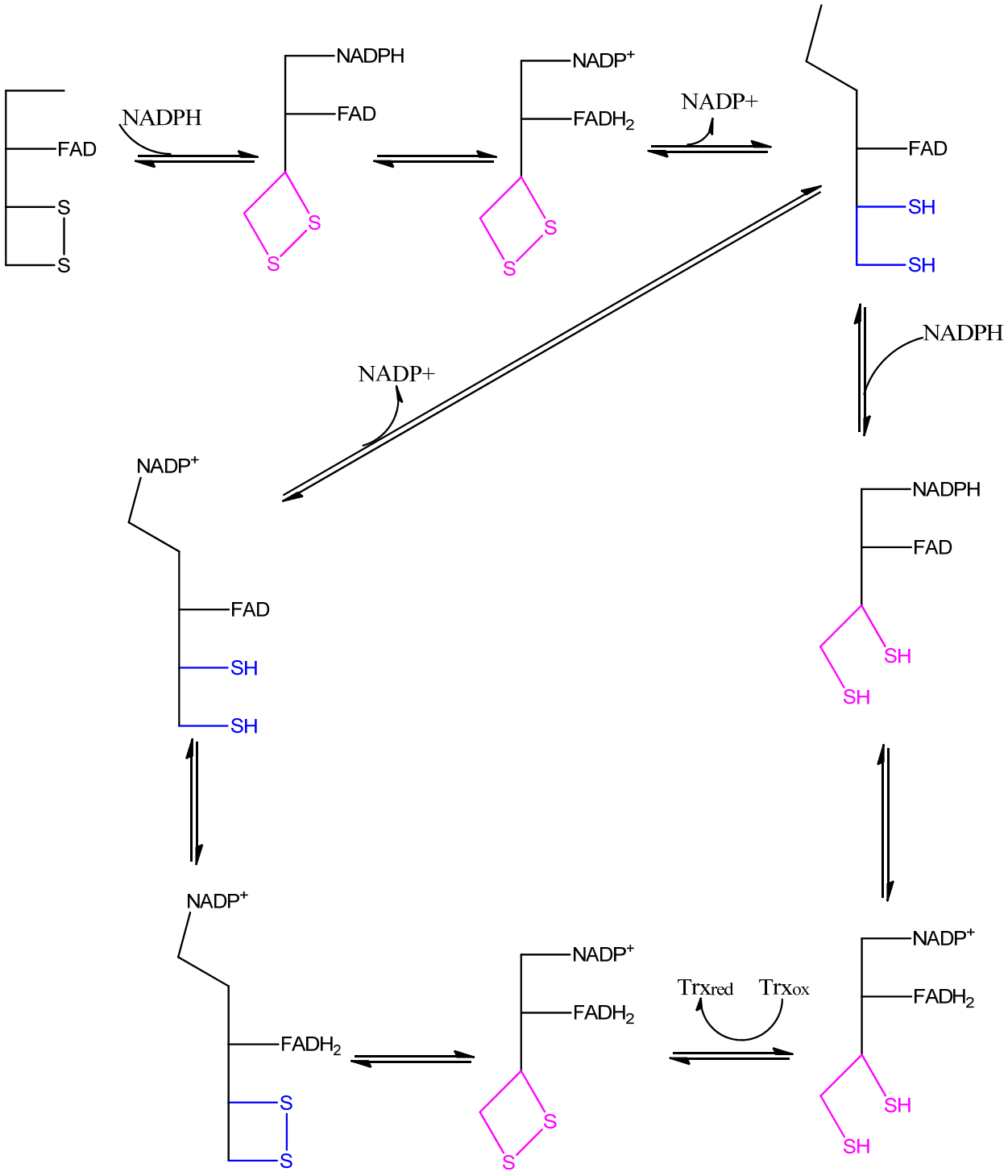


Figure 1.2: Schematic representation of TrxR conformation during catalysis; flavin oxidizing (FO) conformation (Blue) and flavin reducing (FR) conformation (Pink) (Lennon & Williams, 1996). Catalysis of TrxR is accompanied by rotation events and cycling between a two electron and a four electron reduced state of the enzyme.

1.2.2 Structure of Trx

Trx is a relatively small and ubiquitous protein with a molecular weight of about 12 kDa. It contains a highly conserved active site motif WCGPC. The Trx structure consists of the "Trx fold", so-named because it was first identified in Trx (Martin et al, 1993). The Trx fold family includes: glutaredoxin (Bushweller et al, 1994), DsbA (Martin et al, 1993), protein disulfide isomerase (Lundström & Holmgren, 1993), and glutathione peroxidase (Epp et al, 1983). All these proteins are involved in thiol chemistry. The thioredoxin fold consists of a five-stranded beta sheet forming a hydrophobic core surrounded by three or four alpha helices (Figure 1.3). The secondary structural elements of the Trx fold can be subdivided into an N-terminal βαβ motif and a C-terminal ββα motif, connected by a loop of residues that incorporates a third helix (Figure 1.3). The quaternary structure of Trx varies from species to species, however the tertiary structure is highly conserved across various species. Most thioredoxins are monomers in solution, but human thioredoxin forms a homodimer linked by a disulfide (Weichsel et al, 1996).

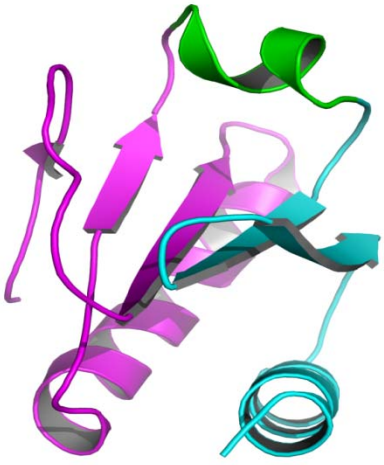


Figure 1.3: Structure of E. coli Trx (PDB code: 1F6M), showing the N-terminal βαβ motif (pink), a C-terminal ββα motif (cyan), and a third helix (green). This figure was generated using PyMol (DeLano, 2002a)

1.2.3 Structure of TrxR

TrxR is a homodimeric enzyme with a highly conserved active site motif CATC. Each monomer consists of an FAD-binding domain and an NADPH-binding domain (Figure 1.4). Both the NADPH-binding and the FAD-binding domains contain variants of the canonical nucleotide binding fold that was first seen in glutathione reductase (Schulz et al, 1978). Each domain contains two β-sheets (one parallel and the other anti-parallel β-sheet) and three α-helices. The two domains are connected by two β-strands and loops, but otherwise separated by a large cleft which is mostly filled with water molecules (Figure 1.4).

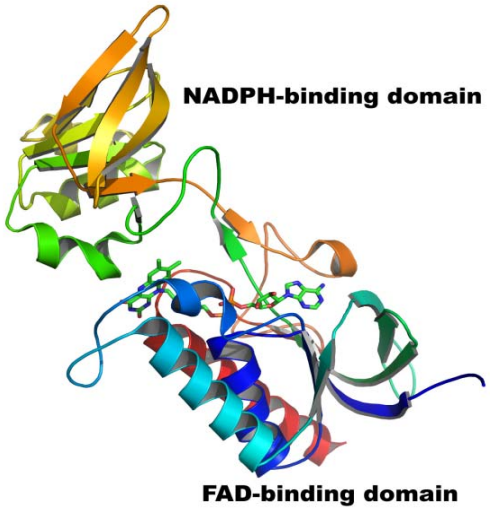


Figure 1.4: Structure of LMW TrxR. Overall structure of the A chain of DR TrxR (PDB code: 2Q7V) in the FO conformation. In this conformation, the FAD-binding domain is too far from the NADPH-binding domain for electron transfer to occur from NADPH to FAD. A 67° rotation by the NADPH-binding domain brings NADPH close to FAD, allowing electron transfer to occur from NADPH to FAD (Kuriyan et al, 1991).

1.2.4 Conformational States of TrxR

The first TrxR structure solved was from *E. coli* (Kuriyan et al, 1991). A surprising finding from the structure was that there was no direct path for the flow of electrons to the active-site disulfide from NADPH via FAD. The NADPH domain was too far (17 Å) from the FAD domain for electron transfer to occur. Moreover, the structure revealed that the active site disulfide is buried and is unavailable for reaction with Trx. This led to the proposal that the enzyme undergoes large conformational changes during catalysis. A 67º rotation places the FAD domain close to the NADPH domain, a favorable position for electron transfer to occur, and exposes the active site thiols for reaction with Trx. A rotation back to the original conformation is then needed for electron transfer to occur from the reduced FAD to the active site disulfide (Figure 1.5). A nomenclature has been suggested (Mulrooney & Williams, 1997) for the two conformations; flavin oxidizing (FO), for the conformation in which the active site disulfide is close to and can oxidize the flavin and flavin reducing (FR) for the conformation in which the active site dithiol is exposed to solvent and NADPH is close to and can reduce the flavin, during catalysis. It is worth emphasizing that the two conformations differ in the domain arrangement rather than in the redox state of the flavin (Figure 1.5). It has been shown that the two conformations are in equilibrium in solution, but the FR conformation is greatly favored in solution (Mulrooney & Williams, 1997). However, all TrxR crystal structures deposited in the PDB are in the FO conformation, thus it can be inferred that the FO conformation predominates the crystalline state. The reason for this is not clear as the "trigger" for rotation is not known (Mulrooney & Williams, 1997).

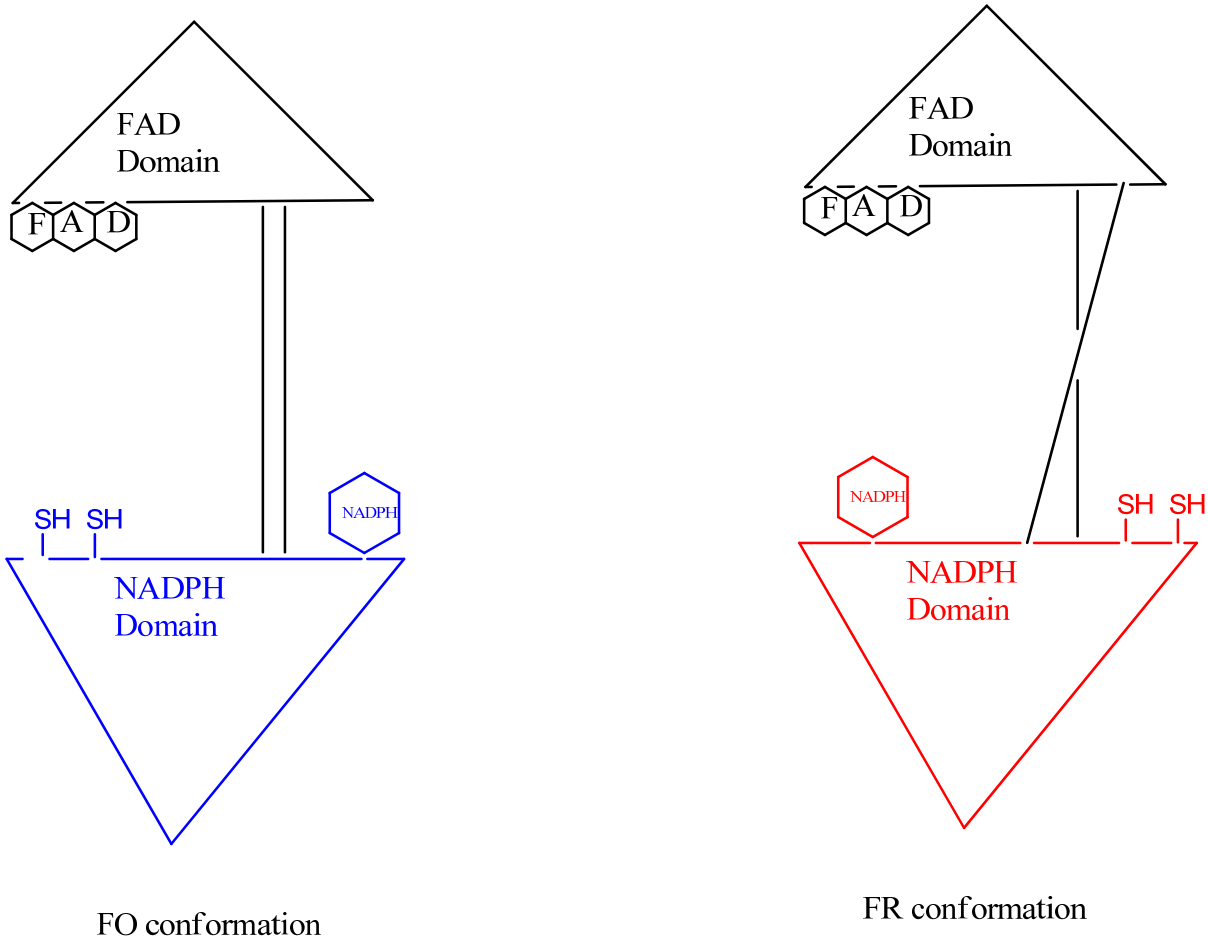


Figure 1.5: Schematic representation of the FO and FR conformations of TrxR. TrxR undergoes a large conformational change during catalysis. From the FO conformation (blue) where the active site disulfide is adjacent to FAD to the FR conformation (red) where NADPH is adjacent to FAD.

Using TrxR and Trx mutants in which one of the active site cysteine residues had been altered to serine in both TrxR and Trx, a complex between TrxR and Trx was successfully formed (Figure 3.24) (Wang et al, 1996). The two proteins were linked by a disulfide bond between their respective remaining single cysteine active-site residues. Based on the proposed FO/FR model, the TrxR-Trx complex is expected to restrict TrxR to the FR conformation, because the FR conformation is the only one in which the active-site thiol is exposed to the protein surface and

can react with Trx. Because of the large size of Trx, it is expected that Trx will stay "trapped" in the FR conformation (Mulrooney & Williams, 1997; Wang et al, 1996). Indeed, when the crystal structure of TrxR-Trx complex was later solved, it was found that TrxR was in the FR conformation when complexed with Trx (Lennon et al, 2000). Therefore, crystal structures of both conformational states have been observed for this enzyme (Figure 1.6): the reduced flavin conformation (FR) conformation (Lennon et al, 2000) and the oxidized flavin (FO) conformation (Waksman et al, 1994).

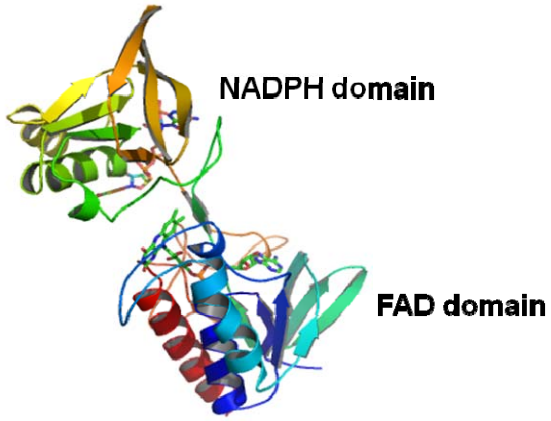


Figure 1.6 A: The structure of the FO conformation of TrxR. In the FO conformation, the NADPH domain is too far from the FAD domain (~17Å) for electron transfer to occur from NADPH to FAD. The figure was generated using HP TrxR structure (PDB code: 2QOK).

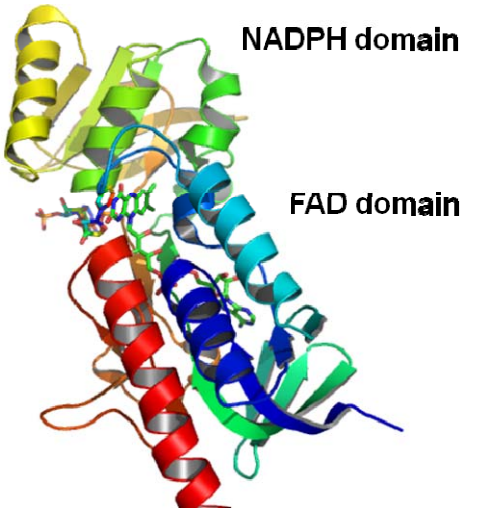


Figure 1.6 B: The structure of FR conformation of TrxR. In the FR conformation, the NADPH domain is close to the FAD domain. Electron transfer can occur from NADPH to FAD. The figure was generated using EC TrxR structure (PDB code: 2F6M).

Fluorescence spectroscopy has been used to study the equilibrium between the two conformations in solution (Mulrooney & Williams, 1997). It was observed that the EC TrxR active site mutant C138S had much less flavin fluorescence (7%) compared to the wild type enzyme (Mulrooney & Williams, 1997). Using this observation Mulrooney and Williams reasoned that flavin fluorescence could be used to study the conformational equilibrium of TrxR in solution. They performed a series of experiments that demonstrated that the two TrxR conformations are in equilibrium in solution, and that the FO conformation is predominant in the C138S mutant and the FR conformation is predominant in the wild type. The large reduction in flavin fluorescence seen in the C138S mutant is presumably because of quenching by a hydrogen bond between the serine at 138 and the N5 of the flavin (Mulrooney & Williams, 1997). This interaction would only be possible in the FO conformation (Figure 1.7; the disulfide is close to the flavin); therefore it can be concluded that the C138S mutant mostly exists in FO conformation in solution. The residual fluorescence seen is probably due to the smaller population of the FR conformation in the C138S mutant.

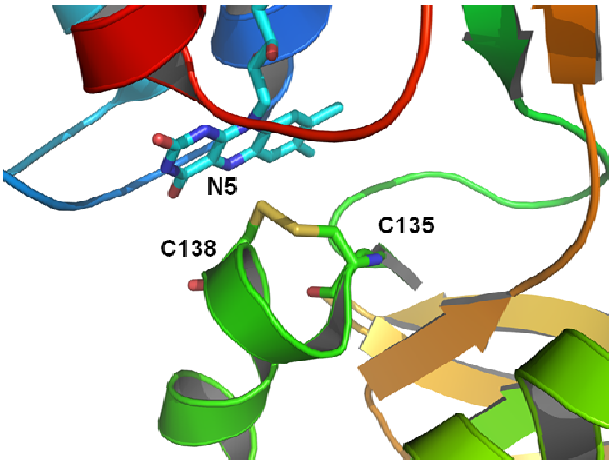


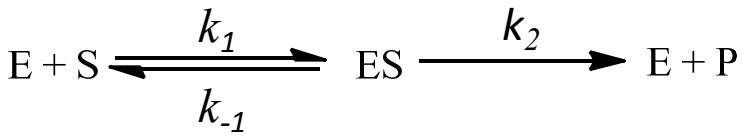
Figure 1.7: Position of the active site disulfide relative to the isoalloxazine ring of FAD. Replacement of the disulfide with a serine residue (EC TrxR C138S mutant) results in quenching by a hydrogen bond between serine 138 and the N5 of the flavin.

Titration of the thiol specific reagent phenyl mercuric acetate (PMA) with the C138S mutant results in a large linear increase of fluorescence which saturates at 1:1 ratio of PMA:C138S, demonstrating that the equilibrium can be altered by binding of the bulk reagents to TrxR active site thiols, and that the change can be monitored by fluorescence. To observe the FR conformation, Mulrooney and Williams (1997), titrated a non-reducing NADPH analog 3-aminopyridine adenine dinucleotide phosphate ($AADP^+$) against the wild type enzyme, PMA treated C138S mutant, and untreated C138S mutant. They observed flavin fluorescence quenching and a shift in $AADP^+$ absorbance maxima in the PMA treated C138S mutant and the wild type enzyme, but not the untreated C138S. These results are consistent with $AADP^+$ analog binding near the FAD in the FR conformation. Therefore it can be concluded that the wild type enzyme exists predominantly in the FR conformation in solution (Mulrooney & Williams, 1997; Van den Berg et al, 2001).

1.3 CHARACTERIZATION OF PROTEIN-PROTEIN INTERACTIONS

1.3.1 Enzyme Kinetics

Enzymes are biological catalysts that increase the rate of chemical reactions without undergoing any permanent changes. They increase reaction rates by providing an alternative lower energy pathway for conversion of substrate to product. Enzymes can increase the rate of the reactions by as much as 10^{17} fold. The basis of their catalytic power is in their ability to generate various localized chemical environments around the same substrate, to bring two substrates to close proximity for reactions to occur, and to form highly reactive covalent intermediates (Price and Stevens, 1999). A simple general scheme for a single substrate enzymatic reaction is shown below:



Usually enzyme reaction rates are measured after a reaction has partially progressed (small percentage) towards the product (initial rates) and, they exhibit saturation at high substrate concentrations. The equation describing enzyme kinetics was derived by making one of two assumptions (apart from the general assumptions that there is no product inhibition, no cooperativity, and the enzyme concentration stays constant throughout the reaction):

(i) The E + S $\rightleftharpoons$ ES equilibrium is only slightly disturbed by breakdown of ES to give product (Michaelis-Menten)

or

(ii) ES remains in a steady state during the reaction (Briggs-Haldane)

According to the steady-state assumption the rate of production of ES (k_1[E][S]) must equal the rate of breakdown (k_{-1}[ES] + k_2[ES])

Therefore: $k_1[E][S] = k_{-1}[ES] + k_2[ES]$

Making [ES] the subject of the formula:

$$[ES] = \frac{k_1[E][S]}{k_{-1} + k_2} = [E][S]K \text{ where } K = \frac{k_1}{k_{-1} + k_2}$$

The fraction of enzyme available as ES, at any concentration can be given as:

$$F = \frac{[ES]}{[E] + [ES]} \quad (1)$$

Therefore:

$$F = \frac{K[E][S]}{[E] + K[E][S]} = \frac{K[E]([S])}{K[E](1/K + [S])}$$

$$\text{where } 1/K = \frac{k_{-1} + k_2}{k_1} = K_m$$

$$F = \frac{[S]}{K_m + [S]} \quad (2)$$

If the total concentration of enzyme is $[E]_0$, then

$$[ES] = F[E]_0 = \frac{[E]_0[S]}{K_m + [S]} \quad (3)$$

The rate of product formation is given by:

$$v = k_2 [ES] = \frac{k_2[E]_0 [S]}{K_m + [S]} \quad (4)$$

Putting $k_2[E]_0 = V_{max}$ then

$$v = \frac{V_{max} [S]}{K_m + [S]} \quad (5)$$ (Michaelis-Menten equation)

Plotting the Michaelis-Menten equation shows that *v* will tend towards a maximum value (*Vmax*) as the value of [S] increases (Figure 1.8). k_2 is known as the k_{cat} or the turnover number; it represents the number of substrate molecules converted into product per unit time. K_m is the Michaelis constant. It may contain more than k_1, k_{-1}, and k_2, since more than one central complex might exist in the course of the reaction. When substrate concentration ([S]) equals the K_m, the Michaelis-Menten equation simplifies to $v = 0.5V_{max}$. For a Michaelis-Menten reaction, k_2 is rate-limiting; therefore $k_2 << k_{-1}$ and K_m reduces to k_{-1}/k_1, which is the dissociation constant for the ES complex. When these conditions hold ($k_2 << k_{-1}$), K_m can be used to describe the affinity of various substrates for the enzyme (Price & Stevens, 1999).

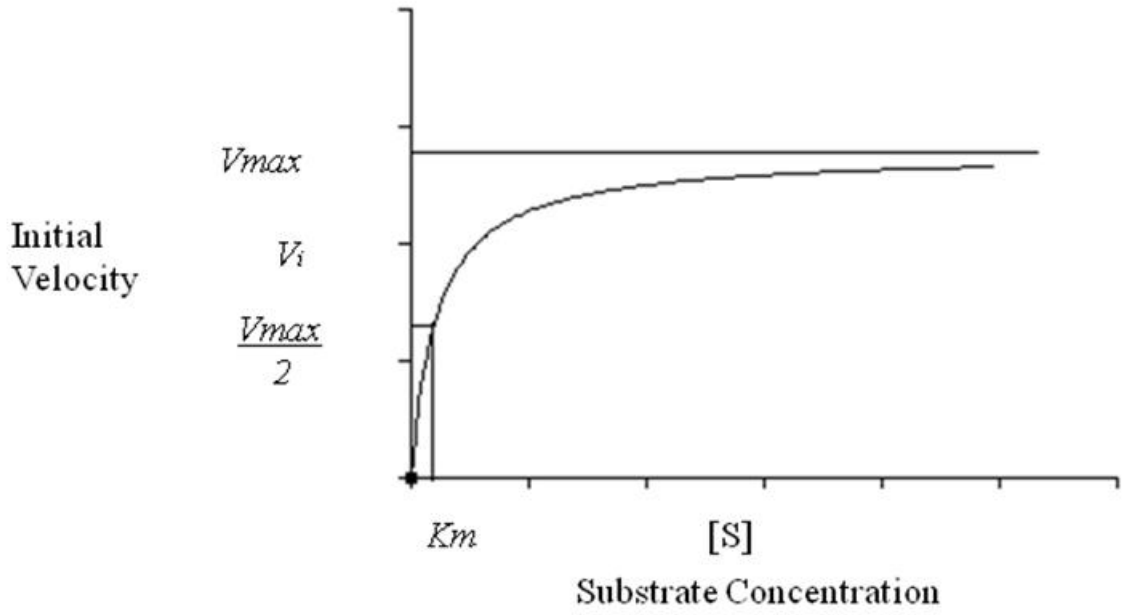


Figure 1.8: A Machaelis-Menten plot. A plot of substrate concentration vs initial velocity for a reaction obeying Michaelis-Menten kinetics

1.3.2 Protein-Protein Docking

Many biological processes depend on specific recognition of one or more proteins by partner proteins (complexation). Examples of protein-protein complexes include: enzyme/inhibitor, antibody/antigen, hormone/hormone receptor, and protein kinase/substrate. Experimental structure determination of protein-protein complexes (primarily by x-ray crystallography and NMR spectroscopy) provides an important tool for characterization of the nature and composition of the interfaces of protein-protein complexes. However, determination of protein complex structures remains difficult, as a result only a few (less than 1 % of the structures deposited in the PDB) hetero-multimeric protein-protein complexes have been solved (Chen & Weng, 2002; Janin, 2007). Computational approaches have previously been used to predict protein-protein complexes formed from two or more free partners. The accuracy of various prediction algorithms is evaluated biennially through community-wide blind prediction

experiments (Critical Assessments of Predicted Interactions (CAPRI)). The degree of success of the computational predictions depends on the amount of conformational change upon complexation, with those complexes showing a small amount of change upon complexation having the highest success (greater than 60% of interface interactions correctly predicted) (Janin, 2007). Two computational approaches are used to model protein-protein interactions (docking and binding simulations). Docking approaches, unlike binding simulations, are not concerned with modeling of binding pathways, but rather focus on the final configuration of the complex (Vakser & Kundrotas, 2008).

Docking approaches involve two elements: a scoring function and a search algorithm. The role of the scoring function is to distinguish between nearly correct and incorrect complexes. The scoring function can be built from empirical knowledge-based data or fundamental physics-based principles. The knowledge-based scoring functions are derived from statistical analysis of biological data bases. These are expressed as potentials that empirically reproduce the fundamental structural properties of proteins. Physics-based methods use energy terms that describe specific interactions between atoms (e.g. covalent bonds, electrostatics, van der Waals and hydrogen bonds) to rank protein-protein complexes (Lee and Lee, 2008). Search algorithms can perform a full solution space search or a guided search that explores only a part of the solution space. Global (full space) searches are typically performed by Fast Fourier Transform methods that perform rigid-body searches i.e. without any flexibility considerations. Some algorithms (e.g. ZDOCK) allow incorporation of known biological information, e.g. predicted interface residues, which simplifies the global search process (Lee and Lee, 2008).

1.4. OVERVIEW OF PROTEIN CRYSTALLOGRAPHY

1.4.1 Protein Crystal Growth

A protein crystal consists of a large number of three-dimensionally arranged repeating units, the unit cell. Each unit cell contains one or more protein molecules whose position and orientation in each unit cell is determined by the space group symmetry. Protein crystal growth can be thought of as a reversible equilibrium process, with molecules entering and leaving the solid phase. At the saturation point, there is no net change in molecules entering and leaving the solid phase. To form crystals it is necessary to achieve supersaturation, where the equilibrium is shifted to favor the solid state. Crystal growth occurs in three stages: nucleation, growth and cessation of growth. In the nucleation stage a "critical nucleus" must be formed; anything smaller will redissolve. The critical nucleus initiates aggregation in an ordered manner leading to the formation and growth of crystals. There has been only limited research done on cessation of growth, as a result, it is not a well-understood process. There are two possible reasons for cessation of growth: a drop in saturation level as crystals form, and/or inhibition of growth by surface defects (McPherson, 1999). A two dimensional solubility phase diagram can be used to visualize the process of crystal growth (Figure 1.9). The solubility curve divides the concentration space into undersaturation and supersaturation regions. The supersaturation region is sub-divided into three zones: metastable zone, nucleation zone, and precipitation zone. Proteins will never crystallize in the undersaturation zone. In the precipitation zone excess protein exists as amorphous precipitate instead of remaining in solution. As the name suggests, nucleation occurs in the nucleation zone; crystal growth can also occur in this zone. The metastable region is ideal for growth of few large crystals without nucleation of new crystals (McPherson, 1998).

Various techniques are available for protein crystallization: vapor diffusion, batch crystallization, and dialysis. In vapor diffusion, the initial reagent concentration in the droplet is less than that of

the well. The well solution pulls water from the droplet in a vapor phase thus increasing reagent concentration in the droplet leading to supersaturation. In contrast, the batch method attains supersaturation state instantaneously when the protein solution and precipitant are mixed. In dialysis, the supersaturated state is achieved by gradually increasing the precipitant concentration using a dialysis membrane (McPherson, 1999).

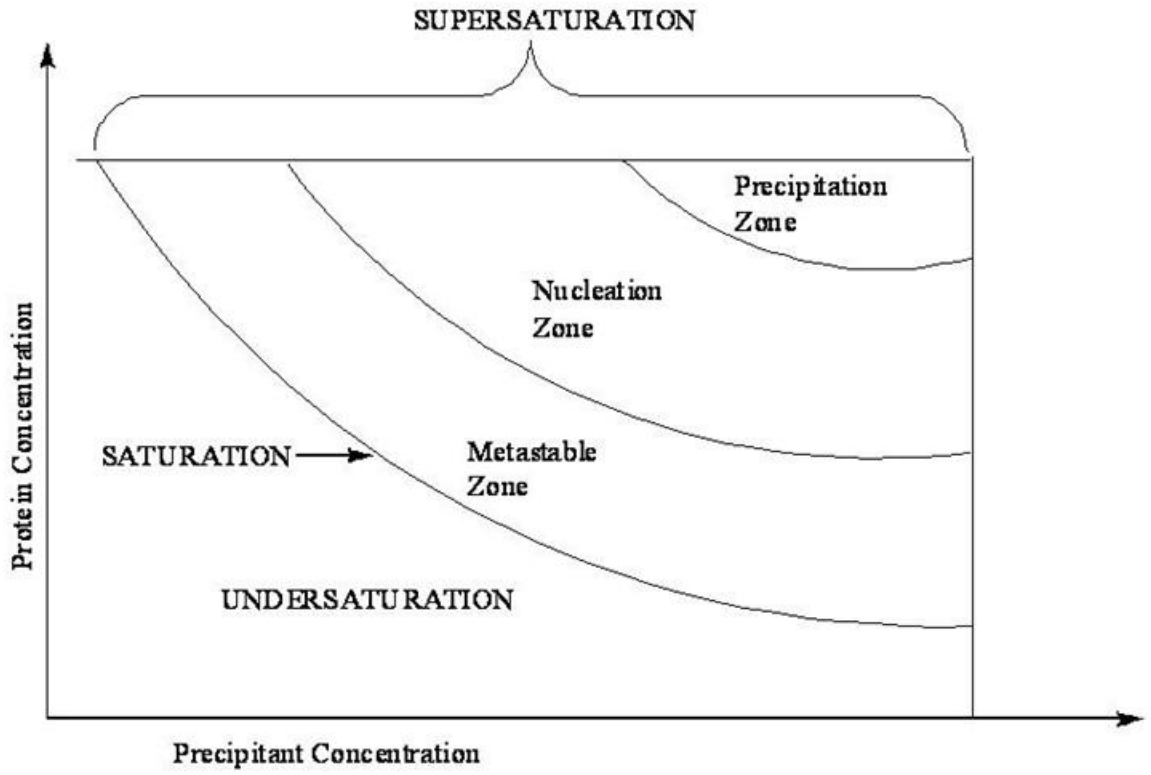


Figure 1.9: Solubility phase diagram for crystallization from a solution.

Obtaining highly diffracting crystals is the biggest bottleneck in protein crystallography. Protein crystals are difficult to grow, because protein molecules are spherical or ellipsoidal in shape, often with irregular surfaces, which do not easily pack without the formation of large holes between molecules. These holes are mainly filled with disordered solvent molecules. Nevertheless, it is possible to grow protein crystals by optimizing various parameters such as

protein purity and concentration, pH and ionic strength of buffer, temperature, precipitants, ligands and additives. The protein itself is perhaps the most important variable in crystallization; therefore, using mutants or homologues can sometimes dramatically improve crystallizability. It is common to screen many different conditions before the ideal set of conditions is found that will produce highly diffracting crystals. The quality of the crystal diffraction is critical to determining the final protein structure (McPherson, 1999).

1.4.2 X-rays and Diffraction Theory

X-rays are a form of electromagnetic radiation with wavelengths of about 0.1-100Å. The wavelength of x-rays is in the same order of magnitude as the bond lengths of atoms in protein molecules (0.5 -1.6 Å) (the reason why x-rays are used to locate atoms in crystals). X-rays can be produced by bombarding a metal target (copper or molybdenum) with high velocity electrons (produced by a heated filament and accelerated by an electric field). An accelerated electron then collides with and displaces an electron from a low-lying orbital in a target metal atom. An electron from a higher orbital fills the resulting vacancy, emitting its excess energy as an x-ray photon. There are three common sources of x-rays used in crystallography: x-ray tubes, rotating anode tubes, and particle storage rings. Of these, particle storage rings which produce synchrotron radiation in the x-ray region are the most powerful x-ray sources (Rhodes, 2000).

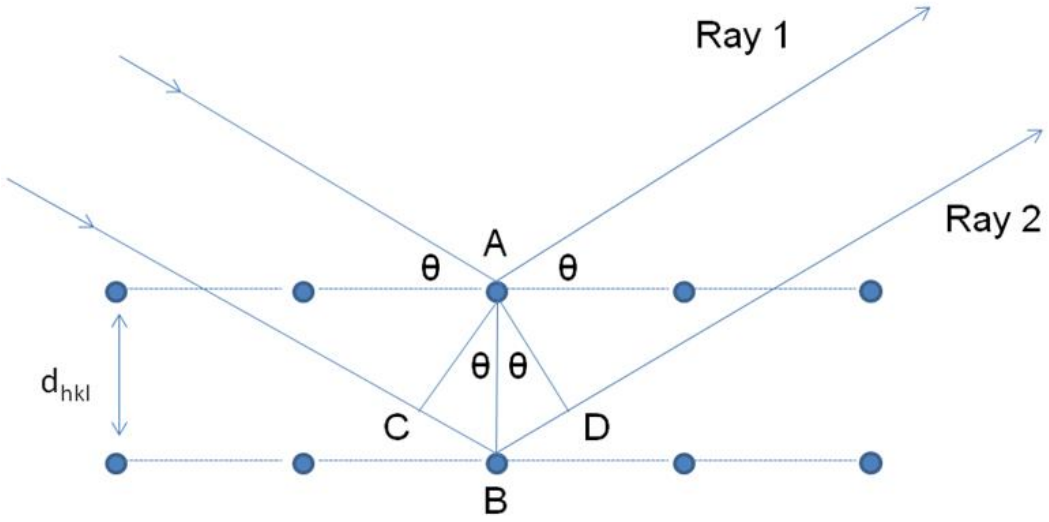


Figure 1.10: Constructive interference conditions: Ray 2 travels a longer distance than ray 1; the additional distance (CB + BD or 2BC) travelled by ray 2 can be given as $2d_{hkl}\sin\theta$ (path difference). If the path difference is equal to the integral number of wavelengths ($n\lambda$) then constructive interference occurs and a strong diffracted beam can be observed.

Diffraction can be defined as scattering (change in the directions and intensities) of waves by objects whose size is approximately the same as the wavelength of the waves. W.L. Braggs (1913) envisioned diffraction of X-rays by crystals to be similar to reflections from sets of regularly spaced parallel planes in the crystal lattice. Scattering occurs in all directions, however measurable reflections (constructive interference) only occur when the path difference between a set of planes is equivalent to integral number of wavelengths i.e. $n\lambda = 2d\sin\theta$ (Figure 1.10). Individual atoms in a molecule can diffract X-rays but are weak scatterers, resulting in weak diffraction that cannot be detected by the instrument. This problem can be solved by analyzing the diffraction pattern of a crystal rather than that of a molecule. Since crystals are composed of identical unit cells, the unit cells will diffract in unison to produce stronger measurable reflections. The diffraction spots measured represent an imaginary lattice (reciprocal lattice) that is inversely related to the real lattice. The reciprocal lattice has the same symmetry as the real lattice (Rupp, 2010; Rhodes, 2000).

It is important to measure intensity for each point in the reciprocal lattice during data collection. To do this, one must first predict which reciprocal lattice points are measurable, given the wavelength and orientation of the incident beam with respect to the crystal. A sphere (Ewald sphere or sphere of reflection) can be constructed such that any lattice point has to lie on its surface to be observed as a reflection i.e. to satisfy Braggs' law. The radius of the sphere of reflection is $1/\lambda$, from the center of the crystal. During data collection, the crystal is rotated so that as many reciprocal lattice points as possible come into contact with the surface of the sphere of reflection. It is possible to move all lattice points within a radius of $2/\lambda$ (from the origin of the reciprocal lattice), to coincide with the surface of the sphere of reflection by rotating the crystal. A sphere with radius $2/\lambda$ is therefore called the limiting sphere corresponding to a resolution of $\lambda/2$ (Rhodes, 2000).

1.4.3 Symmetry

Crystals possess a high degree of internal order (symmetry). The subject of symmetry plays an essential part in all aspects of crystallography. There are seven crystallographic systems based on the relationships between a unit cell's axes (a, b, c) and angles between them (α, β, γ). Combination of the 7 crystal systems with various types of lattice centering, results in 14 lattice types (Bravais lattices). Consequently, there are only fourteen unique ways to arrange identical points in a 3-D space. Points so arranged in the unit cell possess rotary and rotary-inversion operators, leading to 32 unique combinations of symmetry elements (point groups). When translational possibilities are included (screw axis and glide planes), 230 space groups are generated. However, proteins can only crystallize in 65 space groups, because they are chiral. If proteins were to crystallize in some space groups with symmetry operations like rotary-inversion

centres, they will change the asymmetry of their amino acids from L to D (Rupp, 2010; Rhodes, 2000).

The knowledge of crystal symmetry can aid data collection. Due to crystal symmetry, not all reflections in the sphere of reflections are unique; therefore, if the crystal symmetry is known only a fraction of the possible reflections needs to be measured. In addition to all the symmetry elements present in the real space, the reciprocal lattice has a center of inversion which further reduces the number of unique reflections by half (Friedel's law). Crystal symmetry causes some specific reflections to disappear from the diffraction pattern (intensity of zero), due to destructive interference of those reflections. Information obtained from the missing reflections (systematic absences) can be used to derive the space group (Rupp, 2010; Rhodes, 2000).

1.4.4 The Structure Factor Equation

X-rays are scattered when they impinge on a crystal. The resulting diffraction pattern can be recorded by scintillation counters or area detectors. The relative positions and intensities of the reflections on a diffraction pattern contain information about the internal structure of the crystal. All atoms in the unit cell contribute to each reflection of the diffraction pattern as well to angles in between those reflections (because reflections result from constructive interference of all scattered rays that obey Bragg's law), and each reflection represents a unique set of planes. Intensities have to be converted to structure factor amplitudes, before they can be used in the Fourier transformation to get the electron density. Intensities are proportional to the square of structure factor amplitudes.

The structure factor (F_{hkl}) is a complex number which represents the total sum of amplitudes and phases from the entire unit cell contents. It is obtained by adding up the scattering contribution

from each atom in the cell (f_{hkl}) to the reflection *hkl*. The overall structure factor equation of a reflection is thus given by:

$$F_{hkl} = \Sigma f_n e^{i\alpha}$$

Where: i is the imaginary number($\sqrt{-1}$), f_n is the scattering power of the atom and α is the phase angle of the scattered ray relative to the origin of the unit cell. Diffraction of x-rays by atoms is caused by electrons instead of the nucleus of an atom. The structure factor equation is therefore related to electron density and can be described by:

$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z)\, e^{i\alpha}\, dxdydz$$

or

$$F_{hkl} = \int_v \rho(x, y, z)\, e^{i\alpha} dv$$

Where: ρ is the electron density, (x, y, z) are fractional coordinates of an atom in the unit cell and v is the unit-cell volume (Rhodes, 2000).

1.4.5 The Phase Problem

X-rays are electromagnetic waves with phase angles, amplitudes and wavelengths. In order to solve a crystal structure based on reflections recorded on an x-ray film, one needs to know the phase angle, amplitude and wavelength of each reflection. The wavelength of the diffracted beam will be the same as the wavelength of the incident beam and the amplitude can be obtained by finding the square root of intensities measured by the detector. Finding the phase of the diffracted beam is a more difficult problem. Several methods have been developed over the years to overcome the phase problem in protein crystallography: anomalous dispersion, isomorphous replacement, and molecular replacement. For proteins which have homologous structures already solved such as those discussed in this thesis, molecular replacement is generally the method of choice because it does not require further experimental procedures.

1.4.6 Structure Solution by Molecular Replacement

Molecular replacement uses a known protein structure (model structure) to estimate phases of another protein, if there is sufficient sequence identity (usually >40%). The success of this technique is based on the fact that proteins with high sequence identity often have similar three-dimensional structures. To do molecular replacement, the known structure needs to be put in the correct position and orientation in the unit cell of the unknown protein. This process requires two steps: rotation and translation. First, the rotation function puts the known molecule (search model) in the same spatial orientation as the unknown molecule, and then the translation function superposes the correctly-oriented molecules. In both rotation and translation functions, the Patterson functions of the protein molecules are used. It is seen from section 1.2.4, that a Fourier transformation of the structure factors (amplitudes and phases) gives the electron density. What happens when a Fourier transformation of squared amplitudes only is performed? The result is a Patterson map, whose peaks represent vectors between atoms in the unit cell (Drenth, 2007).

Atoms within the same molecule have short vectors which are close to the origin (intra-molecular vectors). If there were no intermolecular vectors, the intra-molecular vectors of the known and unknown protein structures would be the same, except for a rotation difference. Therefore, intra-molecular vectors provide information about the rotational relationship between the known and the unknown molecular structures. There are software packages available that can calculate the intra-molecular vectors for a large number of different orientations using various mathematical procedures: Rossman and Blow procedures, Crowther's fast rotation function, and maximum likelihood techniques. The orientation that has the best overlap between the known

and unknown models of intra-molecular vectors is taken as a possible correct orientation (Drenth, 2007).

Once the correct rotational orientation has been obtained, the correct position of the known model within the unit cell of the unknown molecule is determined by translation. The Patterson functions of atoms from a different molecule within the same unit cell (symmetry-related molecules) are called intermolecular vectors. They are typically longer vectors which provide the translational relationship between the known and unknown protein structures. To locate the correct position, the known molecule is moved through the asymmetric unit of the unknown structure (usually by Fourier methods). The structure factors are then calculated and compared with the observed structure factors, by calculating an *R*-factor or the correlation coefficient, as a function of the molecular position (see Section 1.4.7). The position with the best correlation would be taken as the starting point for model building and refinement (starting phase angle to calculate electron density) (Drenth, 2007).

1.4.7 Model Building, Refinement, and Validation

Structure refinement is an iterative process which involves rebuilding (tweaking) of the model to find a closer agreement between the calculated and observed structure factors. First, the amino acid sequence is changed from that of the search model to the sequence of the protein of interest. The model is then manually adjusted to match the electron density map. The adjusted model is used to estimate new phase angles and recalculate the electron density. The process of manual adjustment and electron density calculation is then repeated several times, until the model is sufficiently refined, as judged by electron density and the crystallographic *R*-factor. Once the model is sufficiently refined, water and other structural elements (e.g. cofactors, ligands) are

included in the model. The crystallographic *R*-factor is a measure of closeness between the model and the real structure. It is calculated by the equation below:

$$R = \frac{\sum||Fo|-|Fc||}{\sum|Fo|} \text{ X } 100$$

Where $|Fo|$ and $|Fc|$ are observed and calculated structure factors respectively. There are several software packages available for crystallographic refinement. Examples include: CNS (Brunger, 2007), REFMAC (Murshudov et al, 1997) and PHENIX (Adams et al, 2010). These programs use one of two methods for refinement: maximum likelihood refinement or least-squares refinement. All refinement programs use empirical restraints or constraints (bond lengths, bond angles, torsion angles, and van der Waals contacts) to arrive at a reasonable structure when refinement is complete.

A well refined high resolution structure (< 2 Å) structure would have an *R*-factor of less than 20% whereas a random acentric structure would have an *R*-factor of 59%. However, it has been found that some incorrectly-built protein structural models can also have *R*-factors of less than 20% due to model bias. A cross-validation index (*R*-free) is used to detect this problem. Before the start of refinement, unique reflections are divided into two groups: a "test set" and a "working set". The test set consists of 5~10 % of the observed reflections and the working set consists of all the remaining reflections. Refinement is then carried out with the working set only, and the test set is used for cross-validation according to the equation above; except that the structure factors, $|Fo|$, of the test set reflections are used instead of the working set reflections (Drenth, 2007).

1.5 MODEL SYSTEMS

1.5.1 *D. radiodurans*

Deinococcus radiodurans (D. radiodurans) is a Gram positive bacterium capable of withstanding exposure to extreme γ-ray and ultraviolet radiation, oxidants, and desiccation. *D. radiodurans* can survive exposure to γ-rays (more than 4000 Gy). In contrast, exposure to less than 200 Gy is lethal to most other bacteria (Cox & Battista, 2005; Daly, 2009). Because of its ability to withstand cellular toxicity caused by oxidants and ionizing radiation, transgenic *D. radiodurans* may be useful in bioremediation of waste sites contaminated with radionuclide contaminants (Brim et al, 2006). As a result *D. radiodurans* has been the subject of intense research aimed at understanding the basis of its ability to survive exposure to extreme conditions, and how it can be used in bioremediation of toxic waste sites (Brim et al, 2006; Cox & Battista, 2005; Dennis et al, 2006). *D. radiodurans'* unusual capabilities have been attributed to a high number of genome copies (Daly et al, 2004), ring-like nucleoid organization (Levin-Zaidman et al, 2003), high manganese content (Daly et al, 2004), and a higher ability to scavenge ROS (Tian et al, 2007). However, the mechanism responsible for its extremophilic nature is not clearly understood (Makarova et al, 2001).

Although ionizing radiation, desiccation and the presence of oxidants can cause different types of damage to the genome, DNA double stranded breaks (DSB) are considered to be the most lethal (Daly, 2009). As a result, efforts to understand the mechanism behind *D. radiodurans'* ability to tolerate extreme conditions have been focused on understanding its ability to prevent or repair DSB (Daly, 2009). Indeed, some experimental evidence suggests that efficient DNA repair is solely responsible for *D. radiodurans'* ability to withstand ionizing radiation. *D. radiodurans'*

sustains the same amount of DNA damage at high radiation doses as other bacteria, but unlike other bacteria its damage is mended within hours (Daly, 2009; Makarova et al, 2001). However, some recent evidence suggests that it is likely that prevention of DNA damage (ROS scavenging) supplements DNA repair to make *D. radiodurans'* ionizing radiation tolerant. It is worth noting that only about 20% of radiation-induced damage to the genome is due to the direct effect of irradiation (the rest is due to radiation-induced ROS), and that cellular extracts of *D. radiodurans* are more effective in scavenging ROS than *E. coli* extracts when subjected to oxidative stress (Tian et al, 2007). Moreover, *D. radiodurans* has higher basal levels of some antioxidant enzymatic systems (catalase and superoxide dismutase) and disruption of superoxide dismutase (sodA) and catalase (katA) genes results in increased sensitivity of *D. radiodurans* to ionizing radiation. In addition, *D. radiodurans* catalase is more resistant to inhibition by substrate H_2O_2, than bovine or *Aspergillus niger* catalase (Kobayashi et al, 2006). Taken together these experimental results suggest a significant contribution of antioxidant systems to *D. radiodurans* ability to withstand extreme ionizing radiation. While the contribution of some antioxidant enzymatic systems to the extremophilic nature of *D. radiodurans* has been extensively studied, the role of the thioredoxin system has not been investigated (Tian et al, 2007). *D. radiodurans* has two thioredoxin genes and a single thioredoxin reductase gene which form part of its defence against oxidative damage. Therefore, as part of the enzymatic systems that protect *D. radiodurans* against ROS damage, characterization of the thioredoxin system may contribute to understanding the extremophilic nature of *D. radiodurans*.

1.5.2 *H. pylori*

Helicobacter pylori (H. pylori) is a microaerophilic, gram-negative bacterium capable of colonizing the mucosal lining of the stomach. This gastric colonization has been associated with peptic ulcer, gastritis, and gastric cancer (Baker et al, 2001a). *H. pylori* avoids the highly acidic environment of the stomach by living within the thick mucus lining protecting the stomach from gastric juice. Its ability to survive in the stomach is enhanced by its motility (Eaton et al, 1992), activity of its urease enzyme (Marshall et al, 1990), and its gastric adherence properties (Zhang et al, 2002). It contains flagella which helps it to move rapidly into the mucosal lining of the stomach in order to avoid long exposure to gastric juices (Eaton et al, 1992). In addition to rapid movement, *H. pylori* generates ammonia (base) through the action of its urease enzyme to protect itself from stomach acids (Marshall et al, 1990). Urease catalyzes the hydrolysis of urea into ammonia and carbon dioxide. Ammonia is then used to neutralize the stomach acid. Although the majority of the colonizing bacteria remain in the mucosal lining of the stomach, some adhere to the gastric epithelial cells. This adherence is facilitated by its adhesins and other virulence factors (Zhang et al, 2002).

Upon colonization, *H. pylori* secretes immunogenic products that recruit macrophages to the site of infection producing a large amount of ROS. Long term exposure to ROS damages gastric tissues, and is thought to be a causative factor in human gastric cancer (Baker et al, 2001a; Correa, 1995; Davies et al, 1994). To protect itself from ROS damage, *H. pylori* has several protective enzymatic systems including catalase, superoxide dismutase, and thioredoxin. However, it lacks the glutathione system which is found in other bacteria (Baker et al, 2001a). Like *D. radiodurans*, *H. pylori* also contains two Trxs (Trx1 and Trx2) and a single TrxR as part

of its Trx system. Both *H. pylori* Trx1 and Trx2 show high sequence identity (51 and 41 % respectively) to *E .coli* Trx (Baker et al, 2001a; Windle et al, 2000). However, Trx2 shows an unusual active site motif (CPDC), instead of the highly conserved active site motif (CGPC) found in other Trxs (Windle et al, 2000).

The *H. pylori* Trx system is somewhat unique among bacterial Trxs. *H. pylori* secretes Trx when subjected to a variety of stresses; the only other known example of secreted Trx is human Trx. It has therefore been postulated that the Trx system plays a role (reduction of disulfides) in colonization of the disulfide-rich mucosal lining of the stomach; a property that may be essential for its viability (Windle et al, 2000). Because of its role in ROS scavenging and possibly in infiltration of the gastric mucin layer, the *H. pylori* Trx system has been mentioned as a potential new antibacterial drug target (Gustafsson et al, 2007; Windle et al, 2000). Therefore characterization of the *H. pylori* Trx system may provide insight on how protein-protein interactions of the *H. pylori* Trx system could be disrupted.

1.6 INDUSTRIAL AND PHARMACEUTICAL APPLICATIONS OF THE THIOREDOXIN SYSTEM

The Trx system has the potential for many pharmaceutical and industrial applications. Exogenous addition of all Trx system components improved the quality of wheat flour by reducing the disulfide bonds of storage proteins (e.g. glutelins and prolamins) making the wheat flour more digestible (Joudrier et al, 2005). Transgenic barley overexpressing Trx was found to have shorter germination periods because of the enhanced activity of starch de-branching enzyme (Gautier et al, 1998; Joudrier et al, 2005). Reduction of disulfides of some food allergens, e.g. from wheat (gladians and glutenins) and milk (β-lactoglobulin) by Trx attenuated their allergenicity (Buchanan et al, 1997; Joudrier et al, 2005). Many venom toxins from snakes,

bees, and scorpions are inactivated by reduction, since their native form contains disulfides (Joudrier et al, 2005). Success of these applications will depend on efficient protein-protein interactions of various Trx system partners. Therefore, the ability to fine-tune the binding specificity of TrxR-Trx and/or Trx-target protein will enhance many of these potential applications of the Trx system. A systematic study of the TrxR-Trx interface can provide new insight on the basis of the binding specificity and may simplify the process of redesigning the Trx system with new specificities and/ or improved efficiencies.

1.7 AIMS AND OBJECTIVES

This work was undertaken with three goals.

1. To determine the affinity of *D. radiodurans* TrxR for *E. coli* and *D. radiodurans* Trxs
2. To determine the redox activity of the *D. radiodurans* Trx system
3. To determine the basis of species-specificity of TrxR-Trx interactions

Previous studies have shown that TrxRs have a higher affinity for cognate Trxs over Trxs from other species (Baker et al, 2001a; Gustafsson et al, 2007). The first objective was to determine if DR TrxR showed higher affinity for its cognate Trx over Trxs from another species (*E. coli),* using spectrophotometric and biophysical approaches.

It has also been previously shown that cell extracts of *D. radiodurans* are more efficient in scavenging ROS than those from *E. coli,* and that the higher ROS scavenging capability of *D. radiodurans* contributes to its extremophilic nature (Daly, 2009; Mattimore & Battista, 1996). The second objective was to determine the redox activity of the *D. radiodurans* Trx system using spectrophotometric assays, and to compare it with that of *E. coli*.

Previous studies have shown that TrxRs have a higher affinity for cognate Trxs than Trxs from other species. The differences in specificity have been attributed to differences in the size of loop

regions of the FAD-binding domains (in eukaryotic LMW TrxR) (Zhang et al, 2009) and differences in the charge and shape of thioredoxin-binding surfaces (in bacterial LMW TrxR) (Gustafsson et al, 2007; Obiero et al, 2010), however no systematic study has been carried out to determine the basis of specificity. The third objective was to determine the basis of species-specificity of TrxR-Trx interactions, and the contribution of specific amino acid residues to TrxR-Trx interface stability and species specificity, using x-ray crystallographic, computational, and biochemical approaches.

CHAPTER TWO: MATERIALS AND METHODS

2.1 PROTEIN OVEREXPRESSION

Conditions for overexpression and induction for various proteins are shown on Table 2.1 below. After overexpression and induction the cells were harvested by centrifugation (20 min, 8000 g, 277 K) and the pellets stored at -80 ºC.

Plasmid	Host cell strain	Media	Antibiotics (μg/L)	Induction OD_{600}	[IPTG] μM	Harvesting time after induction	Induction temperature
pPROK1/Trx1	BL21Gold	LB	Ampicillin (100)	0.9	0.4	16 hours	37ºC
pPROK1/Trx2	BL21Gold	LB	Ampicillin (100)	0.6	0.4	3 hours	37ºC
pPROK1/TrxR	Rossetta2	LB	Ampicillin (100)	0.9	0.4	16 hours	37ºC

Table 2.1A: Overexpression and induction conditions for *H. pylori* proteins

Plasmid	Host cell strain	Media	Antibiotics (μg/L)	Induction OD_{600}	[IPTG] μM	Harvesting time after induction	Induction temperature
pHISTEV/Trx1	Rossetta2	LB	Kanamycin (50)	0.6	0.25	4 hours	30ºC
pHISTEV/TrxR	Rossetta2	LB	Kanamycin (50)	0.6	0.4	16 hours	30ºC

Table 2.1B: Overexpression and induction conditions for *D. radiodurans* proteins

Plasmid	Host cell strain	Media	Antibiotics (μg/L)	Induction OD_{600}	[IPTG] μM	Harvesting time after induction	Induction temperature
pHISTEV/Trx1	Rossetta2	LB	Kanamycin (50)	0.6	0.4	4 hours	37ºC
pHISTEV/TrxR	Rossetta2	LB	Ampicilin (100)	0.6	0.4	16 hours	30ºC

Table 2.1C: Overexpression and induction conditions for *E. coli* proteins

2.2 PROTEIN PURIFICATION

2.2.1 Purification of HP Trx1

Helicobacter pylori thioredoxin-1 (HP Trx1) was purified as previously described (Baker et al, 2001a; Pervin, 2006). Briefly the cells were induced and harvested by centrifugation (Table 2.1A). The crude cell extracts were suspended in 25 mM potassium phosphate buffer (pH 7.0) containing 1 mM ethylenediamine tetra-acetic acid (EDTA) and stirred at 4°C for 30 min in the presence of DNase and lysozyme (2.5% w/v). After sonication and centrifugation, the supernatant was subjected to heat denaturation at 70°C for 4 min to remove contaminating proteins. Denatured proteins were removed via centrifugation and the supernatant was dialyzed overnight in three changes of 5 mM potassium phosphate buffer (pH 7.0). The dialyzed protein was filtered using an acrodisc syringe filter and loaded to a diethylaminoethyl-cellulose (DEAE-cellulose) column which was pre-equilibrated with 10 mM potassium phosphate buffer (pH 7.0). A linear gradient of 10 to 100 mM potassium phosphate was applied to elute the protein of interest. The resulting fractions were analyzed on a sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE) gel and then concentrated to 5 mL before loading onto a gel filtration column. Twenty five mM potassium phosphate buffer with 1 mM EDTA was used as

equilibration buffer and the protein was eluted with 200 mM NaCl. The purity of the protein was assessed by SDS-PAGE and NaCl was removed by dialysis against 25 mM potassium phosphate buffer with 1 mM EDTA. The pure protein was concentrated to an appropriate volume, before measuring the concentration using Bradford assay (1976) and storage as aliquots at -80°C.

2.2.2 Purification of HP Trx2

After centrifugation *Helicobacter pylori* thioredoxin-2 (HP Trx2) crude extracts were treated as described for HP Trx1, except for the following modifications. Contaminating proteins were removed by a 30% ammonium sulfate precipitation and a 4 minute 50°C heat denaturation. The resulting supernatant was dialyzed against 5 mM potassium phosphate buffer pH 7.0. The dialyzed protein was then applied to the anion exchange (DEAE-cellulose, Whatman) column pre-equilibrated with 5 mM potassium phosphate buffer, pH 7.0 and the flow-through fraction was collected. The flow-through fraction was then concentrated and loaded onto a cation exchange column (carboxymethyl cellulose, Whatman) pre-equilibrated with 5 mM potassium phosphate buffer, pH 7.0. The protein was eluted with a linear gradient of 5 to 60 mM potassium phosphate, pH 7.0. SDS-PAGE was again used to assess purity and the protein was stored at -80°C after measuring its concentration by the Bradford assay.

2.2.3Purification of HP TrxR

Purification of *Helicobacter pylori* thioredoxin reductase (HP TrxR) has also been previously described (Baker et al, 2001a; Pervin, 2006). After induction, the cells were harvested by centrifugation. The cell pellets were then resuspended in sonication buffer (50 mM potassium phosphate buffer pH 7.0, 1 mM 4-(2- Aminoethyl) benzenesulfonyl fluoride (AEBSF), 20 mg ml^{-1} DNase and 20 mg ml^{-1} lysozyme), and incubated at 4°C for 30 min. After sonication and

centrifugation, the supernatant was subjected to heat denaturation 4 min at 50ºC, and 20% ammonium sulfate treatment to remove contaminating proteins. An 80% ammonium sulfate treatment was then applied to the resulting supernatant to precipitate TrxR. After the 80% ammonium sulfate treatment, the protein was resuspended and dialyzed (10 mM potassium phosphate buffer pH 7.0) before loading onto the DEAE-cellulose column and elution with a linear gradient with NaCl (20 mM potassium phosphate buffer pH 7.0 to 500mM NaCl, 20 mM potassium phosphate buffer pH 7.0). Fractions containing TrxR, as assessed by the ratio of A_{280}/A_{450} were pooled, concentrated and loaded onto an Affinity Gel Blue column (BIORAD) preequilibrated with a 25mM potassium phosphate buffer, pH 7.0. The protein was then eluted with a linear gradient of potassium phosphate buffer (25 mM potassium phosphate buffer pH 7.0 to 0.3 M NaCl in 600 mM potassium phosphate pH, 7.0). The purity of the final purified TrxR fractions was assessed by SDS-PAGE. Those fractions that were judged to be of high purity were pooled and dialyzed overnight against 50 mM Tris–HCl pH 8.0 buffer.

2.2.4 Purification of DR TrxR and Mutants

Deinococcus radiodurans thioredoxin reductase (DR TrxR) was purified as previously described (Obiero et al, 2006). The frozen DR TrxR cell pellets were thawed and resuspended in sonication buffer (50 mM Tris–HCl, pH 8.0, 1 mM AEBSF, 20 mg ml^{-1} DNase and 20 mg ml^{-1} lysozyme). The thawed cells were mechanically disrupted by sonication and cell debris was removed by centrifugation. The resulting supernatant was loaded onto a POROS MC50 metal-chelation column (Applied Biosystems, USA) pre-equilibrated with a buffer containing 5 mM imidazole, 0.5 M NaCl, 50 mM Tris–HCl, pH 8.0. The column was washed with a 60 mM imidazole, 0.5 M NaCl, 50 mM Tris–HCl pH 8.0 buffer to remove non-specific binding and the protein eluted with a 0–1 M imidazole gradient. Fractions (10 ml) were collected and the purity of the protein was

checked on Coomassie-stained SDS–PAGE. Those fractions showing a high level of purity were pooled and dialyzed against crystallization buffer (50 mM Tris–HCl, pH 8.0). All DR TrxR mutants were purified as described for the wild type protein.

2.2.5 Purification of DR Trx1

DR Trx1 was purified as previously described (Obiero et al, 2010). The frozen DR Trx1 cell pellets were thawed on ice and resuspended in 5 mM EDTA, 50 mM Tris-HCl at pH 8.0, 0.5% Triton X-100, 1 mM AEBSF, 20 mg ml^{-1} DNase, and 20 mg ml^{-1} lysozyme. The thawed cells were mechanically disrupted by sonication and centrifuged (20 min, 8,000 x *g*, 4°C). The protein was found in inclusion bodies by SDS-PAGE analysis of the resulting pellet and supernatant. It was then solubilized from the pellet by resuspending the pellet in 5 mM EDTA, 50 mM Tris-HCl at pH 8.0 and 0.5% DOC, followed by sonication and centrifugation (26,500 x *g*). The resulting supernatant and pellet were separated by decanting. The solubilization procedure was repeated for the pellet. The two supernatants were pooled and loaded onto a POROS MC50 metal chelation column (Applied Biosystems, USA) preequilibrated with a buffer containing 1 mM imidazole, 0.5 M NaCl, and 50 mM Tris-HCl at pH 8.0. The column was washed with the same buffer to remove nonspecific binding, and the protein was eluted with a 0 to 0.5 M imidazole gradient. Peak fractions (5 ml) were collected, and the purity of the protein was checked using Coomassie blue-stained SDS-PAGE. Those fractions showing high levels of purity were pooled and concentrated. The concentrated protein was then dialyzed against 50 mM potassium phosphate buffer at pH 8.

2.2.6 Purification of EC TrxR

EC TrxR was purified using an HQ 20 anion-exchange column (Applied Biosystems, USA) and a Cibacron Blue column. The HQ 20 anion-exchange column was preequilibrated with 25 mM potassium phosphate buffer, pH 8.0. The protein was eluted with a linear gradient of 0 to 500 mM NaCl. The eluted fractions were analyzed by SDS-PAGE. Then the fractions that contained the EC TrxR were applied to an Affinity Gel Blue column preequilibrated with a 25 mM potassium phosphate buffer, pH 7.0. The protein was eluted with a linear gradient of 25 mM potassium phosphate buffer to 0.3 M NaCl in 600 mM potassium phosphate pH 7.0. Peak fractions (5 ml) were collected, and the purity of the protein was checked by SDS-PAGE analysis. Those fractions showing high levels of purity were pooled and concentrated. The concentrated protein was then dialyzed against 50 mM potassium phosphate buffer at pH 8.0.

2.2.7 Purification of EC Trx

EC Trx was purified as described for DR TrxR (Section 2.2.4)

2.3 DETERMINATION OF PROTEIN PURITY

The purity of various protein samples was checked using SDS-PAGE. In this technique, separation based on molecular weights is obtained by the sieving properties of different percentage gels, because of the uniform shape and charge of the protein conferred by SDS and reducing agents. 10% and 15% gels were used to assess the purity level of proteins. These gels were prepared according to standard recipes (Sambrook & Russell, 2001).

2.4 ESTIMATION OF PROTEIN CONCENTRATION

The concentration of the various proteins was estimated using the Bradford assay (Bradford, 1976). First a standard calibration curve was constructed using known concentrations of bovine serum albumin (BSA). The standard samples and unknown protein samples were prepared according to Table 2.3. The Bradford reagent and BSA were purchased from Sigma-Aldrich Canada Ltd. The samples and standards were incubated for 20-30 min at room temperature before measuring the absorbance at 595 nm by UV/VIS spectrophotometry.

Bradford Standard (μg/mL)	0.1 μg/mL BSA (μL)	dH_2O (μL)	Bradford Reagent (μL)
0	0	500	500
1	10	490	500
2	20	480	500
3	30	470	500
5	50	450	500
10	100	400	500
Unknown sample			
1 or 2 μL	0	499/498	500

Table 2.2: Preparation of Bradford standards and unknown sample

2.5 PREPARATION OF MUTANTS

All site-directed mutagenesis experiments were performed using the QuickChange site-directed mutagenesis kit from Stratagene (USA). With the exception of the extension time (which was increased from 1 min/kilobase to 2 min/kilobase), all the PCR parameters were maintained according to the instructions on the kit. The polymerase chain reaction (PCR) reaction mixture contained: 50 ng DNA template, 1X PCR reaction buffer, 125 ng forward primer, 125 ng reverse primer and 2.5 units of *Pfu* Turbo DNA polymerase. The primers used were designed with help of a program (MutPrimer) recommended by the kit. The list of primers is shown in Table 2.3. Following temperature cyling, the reaction was cooled to 25ºC and digested with 10 units of *DpnI* at 37 ºC for 1 hour, to cleave parental methylated DNA. Fifty µL of competent NovaBlue cells were then transformed with 1 µL of the DpnI-digested reaction mixture and added to 250 µL LB media at room temperature. The cells were then grown at 37ºC for 1 hour with shaking at 250 rpm, and the whole transformation mixture was plated on LB agar plates (with appropriate antibiotic). The plates were then incubated overnight at 37ºC. A few colonies from each plate were used to inoculate 4 mL of LB media and incubated at 37ºC overnight with shaking at 250 rpm. After the overnight incubation, the plasmid was isolated using the QIAprep spin miniprep kit and the presence of the desired mutation was confirmed by DNA sequencing (Figure 2.1).

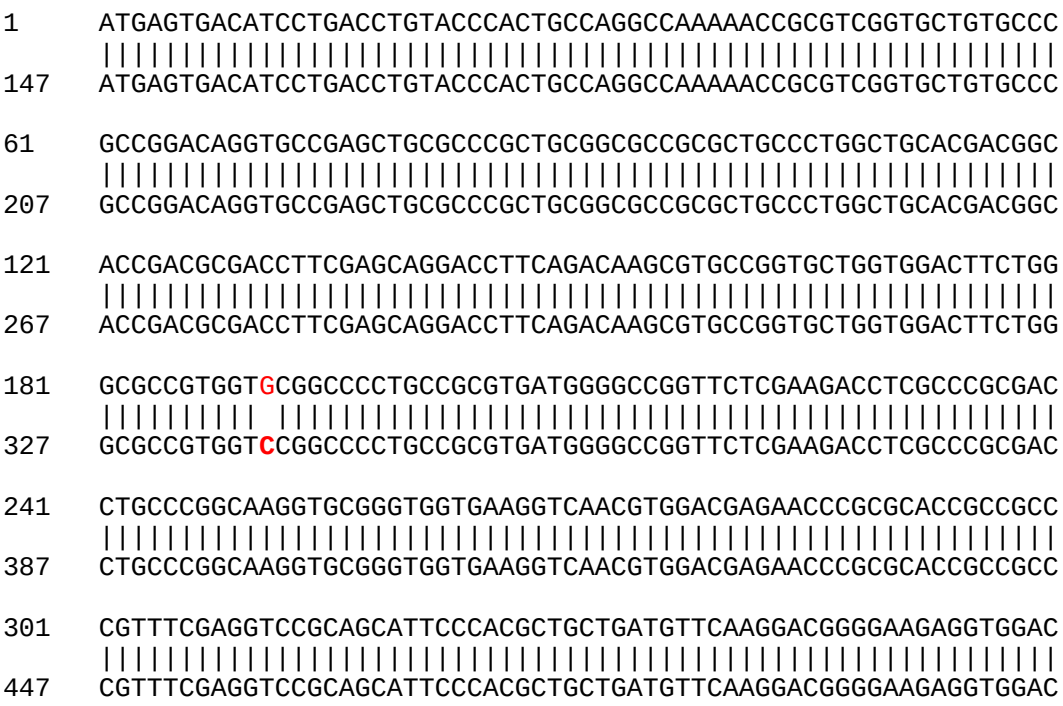


Figure 2.1: Confirmation of mutagenesis by DNA sequence alignment. The mutant sequence (C64S DR Trx1) is shown as the bottom line and parental sequence is shown as the top line. The figure was generated by DNAMAN (Lynnon Corp., USA).

Mutant	Protein	Primer
C133S	*H. pylori* TrxR	Forward primer: 5’ GGTAAAGGCGTTAGCACTAGCGCGACATGC 3’ Reverse primer: 5’ GCATGTCGCGCTAGTGCTAACGCCTTTACC 3’
C30S	*H. pylori* Trx1	Forward primer: 5’ GGCGCCATGGAGTGGGCCTTGTAAGATG 3’ Reverse Primer: 5’ CATCTTACAAGGCCCACTCCATGGCGCC 3’
C142S	*D. radiodurans* TrxR	Forward primer: 5’ CAAAGGCGTGAGCACCAGCGCCACCTGC 3’ Reverse primer: 5’ GCAGGTGGCGCTGGTGCTCACGCCTTTG 3’
C145S	*D. radiodurans* TrxR	Forward primer: 5’ GCACCTGCGCCACCAGCGACGGCTTTT 3’ Reverse primer: 5’ AAAAGCCGTCGCTGGTGGCGCAGGTGC 3’
K137R	*D. radiodurans* TrxR	Forward 5’ GACAACTTCTGGGGC**AGA**GGCGTGAGC 3’ Reverse 5’ GCTCACGCCTCTGCCCCAGAAGTTGTC 3’
K137A	*D. radiodurans* TrxR	Forward 5’ ACAACTTCTGGGGC**GCA**GGCGTGAGCACC 3’ Reverse 5’ GGTGCTCACGCCTGCGCCCCAGAAGTTGT 3’
F148A	*D. radiodurans* TrxR	Forward 5’ CCACCTGCGACGGC**GCT**TTCTATAAGGGCAAGA 3’ Reverse 5’ TCTTGCCCTTATAGAAAGCGCCGTCGCAGGTGG 3’
F149A	*D. radiodurans* TrxR	Forward 5’ CGACGGCTTT**GCC**TATAAGGGCAAGAAAGTCGTGG 3’ Reverse 5’ CCACGACTTTCTTGCCCTTATAGGCAAAGCCGTCG 3’
M84A	*D. radiodurans* TrxR	Forward 5’ GCCAAAGTGGAGGCGGACGAGGTGCAGGGC 3’ Reverse 5’ GCCCTGCACCTCGTCCGCCTCCACTTTGGC 3’
M84F	*D. radiodurans* TrxR	Forward 5’ CCAAAGTGGAGTTCGACGAGGTGCAGGGCGTGC 3’ Reverse 5’ GCACGCCCTGCACCTCGTCGAACTCCACTTTGG 3’
C64S	*D. radiodurans* Trx1	Forward 5’ CTGGGCGCCGTGGTCCGGCCCCTGC 3’ Reverse 5’ GCAGGGGCCGGACCACGGCGCCCAG 3’
C67S	*D. radiodurans* Trx1	Forward 5’ CGTGGTGCGGCCCCTCCCGCGTGATGG 3’ Reverse 5’ CCATCACGCGGGAGGGGCCGCACCACG 3’

Table 2.3: List of primers used for site-directed mutagenesis

Number of repeats	Step	Temperature (ºC)	Time
1	1	95 (denaturation)	30 seconds
16	2	95 (denaturation)	30 seconds
	3	55 (annealing)	1 minute
	4	72 (extension)	14 minutes (2 min/kb)
1	**5**	25 (cooling)	2 minutes

Table 2.4: PCR conditions for mutagenesis

2.6 ACTIVITY ASSAYS

2.6.1 Insulin Reduction Assay

The insulin reduction assay monitors the consumption of NADPH by measuring the decrease in absorbance at 340 nm. The assay mixture (1 mL) contained 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (pH 7.4), 2 mM EDTA, 30 μM bovine insulin (Sigma), 0.1 mM NADPH (Calbiochem), 0.1 μM various TrxRs and various Trxs (0.5 μM -125 μM). All the measurements were carried out at room temperature using a Cary 50 spectrophotometer (Varian). The reaction was initiated by adding Trx and the resulting change in absorbance was recorded. The data were fit to a straight line over a ~1 min range to calculate the change in absorbance/min. Initial velocities were calculated as μM of NADPH oxidized/min in accordance with the relationship $V_o = \Delta A_{340}/0.0062$ (Holmgren & Morgan, 1976). Kinetic parameters were obtained from the plot of initial velocity versus substrate concentration, i.e. the Michaelis-Menten plot.

2.6.2 DTNB Reduction Assay

In this assay 5, 5'- Dithio-bis (2-nitrobenzoic acid) (DTNB) acts as the final recipient of reducing equivalents. The reaction mixture contained 20 mM HEPES (pH 7.4), 2 mM EDTA, 0.24 mM NADPH, 0.1 μM TrxR (various), 1.26 mM DTNB and (0.1-20 μM) Trx (various) in a final volume of 1 mL. The reaction was initiated by adding DTNB and monitored by the increase in absorbance at 412 nm in a Cary 50 spectrophotometer (Varian). Initial velocities were calculated as μM of DTNB reduced/min in accordance with the relationship $V_o = \Delta A_{412}/0.0136$ (Luthman & Holmgren, 1982; Navarro et al, 1991). Kinetic parameters were also obtained from the plot of initial velocity versus substrate concentration.

2.7 FLUORESCENCE MEASUREMENTS

The fluorescence measurements were performed using a Photon Technology International (PTI) Fluorescence Master Systems fluorimeter and the fluorescence spectra recorded with Felix32 software. Purified samples of wild type EC TrxR and EC TrxR C145S mutant were dialyzed in 50 mM K_2HP04 pH 7.5 and then concentrated to 4 mg/mL before doing the fluorescence measurements. The excitation wavelength was set at 455 nm and the emission wavelength was set at 515 nm. Instrumental parameters were: configuration was set to digital-double-double for both excitation and emission scans; slit width of 2.00 nm; hardware integration time of 3 seconds; data collection time of 1 second; average set to 1; number of points for one average was set to 5; lamp power supply was set at 70 watts; step size was set to 1 nm; and voltage of the photon-multiplier was set at 1101V. Data is collected as a voltage signal which is proportional to fluorescence intensity (number of photon/second) with arbitrary units.

2.8 THIOREDOXIN REDUCTASE-THIOREDOXIN COMPLEX FORMATION

Aliquots of a Trx active site cysteine mutant (HP Trx C28S) were treated with dithiothreitol (DTT) by incubation with a ~20 fold excess of fresh DTT at room temperature for 30 min to assure that the remaining thiol was fully reduced. The reaction mixture was concentrated with Amicon filtration units to a volume of about 3 mL and mixed with fresh buffer to a final volume of about 15 mL. The filtration and concentration process was repeated several times (~5 times), until there was no DTT detected in the filtrate (DTT can be detected by reacting the filtrate with a DTNB solution and observing if there is any increase of absorbance at 412 nm). The concentration of the final sample was determined by the Bradford assay. The desalted Trx mutants were then reacted with an ~80-fold excess of DTNB and monitored at 412 nm for the completion of the reaction. DTNB was also removed by several rounds of filtration and concentration. DTNB removal was determined to be complete when there was no more characteristic yellow color in the filtrate. For the *D. radiodurans* active site cysteine Trx mutants, dialysis was used for the desalting process, because the protein easily precipitated upon concentration. A sample of the desalted mixed disulfide Trx-TNB was used directly in the reaction with TrxR active site cysteine mutant (DR TrxR C142S mutant). The TrxR mutant was also pre-treated with DTT (to make sure the remaining active site thiol was also in the reduced state) and desalted in a similar manner to the Trx mutant, before it was reacted with Trx-TNB.

2.9 CRYSTALLIZATION AND CRYSTAL HARVESTING

2.9.1 HP TrxR

HP TrxR was concentrated to 10 mg mL^{-1} (in 20mM Tris buffer pH 8.0) and used for broad screen crystallization trials with sitting drop vapour diffusion. Two conditions (2.2 M ammonium sulfate, 0.2 M lithium nitrate/2.2 M ammonium sulfate, 0.2 M ammonium nitrate) from the ammonium sulfate broad screen yielded crystals. Grid screening and additive screening were performed based on these results to find the optimal crystal growth conditions at 20.0 ^{0}C. The crystals also grew reprodubibly at 4.0 ^{0}C.

2.9.2 DR TrxR

The purified DR TrxR was also concentrated to 10 mg mL^{-1} (in 20mM Tris buffer pH 8.0) and used for crystallization trials at 20.0 ^{0}C. Initial screening was performed by the sitting-drop vapour-diffusion method using broad screens from Qiagen. A 4 mL drop consisting of 2 mL well buffer and 2 mL protein solution was equilibrated against 100 mL reservoir buffer.

2.10 DATA COLLECTION AND PROCESSING

HP TrxR crystals were soaked in cryoprotectant solution for a few seconds prior to flash cooling in a nitrogen cold stream. The cryoprotectant solution contained 2% glycerol, 3.5 M ammonium sulfate and 0.2 M ammonium nitrate. X-ray diffraction data were collected at 100 K on beamline 08-1D.1 using Mar225CCD detector at the Canadian Light Source (Table 3.2). The crystal-to-detector distance was set to 180 mm with an oscillation range of 0.5º and a total of 360 images (1 s exposure time for each image) were collected that covered a total oscillation range of 180º. The images were integrated and scaled using XDS/XSCALE (Kabsch, 1993).

DR TrxR diffraction data were collected using an ADSC Q315 detector on beamline 14-BM-C at BioCARS, Advanced Photon Source (APS), Chicago, IL, USA. The crystal-to-detector distance was maintained at 275 mm with an oscillation range per image of 0.5º, covering a total oscillation range of 90º. The resulting intensity data were indexed and integrated using MOSFLM (Leslie, 1992), scaled and merged using SCALA (Collaborative Computational Project, Number 4, 1994) and converted to structure factors using TRUNCATE (Collaborative Computational Project, Number 4, 1994).

2.11 HOMOLOGY MODELING

Models of DR Trx1 and the FR conformation of DR TrxR were generated by MODELLER 9v1 (with default parameters). Coordinates of the EC Trx crystal structure (PDB code: 2TRX) and EC TrxR crystal structure (FR conformation) (PDB code: 1F6M) were used as templates to build the DR Trx1 model and the DR TrxR model (FR conformation) respectively. Sequence identities are 34% (Trx) and 38% (TrxR) respectively. Side chains of the DR TrxR model (FR conformation) were manually adjusted to match those of the DR TrxR structure (PDB code: 2Q7V) using COOT. Since DR Trx1 has been suggested to be more similar to a classic Trx2 (Obiero et al, 2010) than Trx1, a second model using *R. capsulatus* Trx2 (31% sequence identity) was constructed for comparison purposes (PDB code: 2PPT). The overall quality of the three models was checked by PROCHECK (Laskowski et al, 1993).

2.12 PROTEIN-PROTEIN DOCKING

A number of algorithms have been developed for initial stage of docking (rigid-body docking); including ZDOCK, HEX, FTDock, DOT, DARWIN, and BiGGER (Chen et al, 2003). Most of these algorithms search for the binding modes based on shape complementarity and electrostatics (Chen et al, 2003). Rigid-body docking simulations were done using the ZDOCK server (http://zdock.bu.edu/) with constraints based on the EC TrxR/Trx complex structure (Trx-binding residues were maintained at the interface). The three models (TrxR, Trx1, and Trx2) were used to generate two DR TrxR-Trx complexes. Interacting partners were first placed at a reasonable binding orientation using PYMOL based on the *E. coli* TrxR-Trx complex structure, and then ZDOCK was used to find the final binding mode. The top complex models (best ZDOCK score) were compared to the EC TrxR/Trx complex structure for correctness and then used for shape and complementarity evaluation. Various complexes in Table 3.4 were generated in a similar manner.

2.13 CALCULATION OF INTERACTION ENERGIES

The DR TrxR-Trx complex was minimized using SYBL8.0 implementation of the AMBER7 F99 force fields. A convergence criterion of a minimum energy change of 0.05 $kcal^{-1}.mol^{-1}.Å^{-1}$ was used. The dielectric model consisted of a distance-dependent dielectric constant with the distance cutoff set to 8Å and an ε value of 4.0. AMBER force fields account for forces from van der Waals interactions, bond stretching, bond angles, dihedral (torsion) angles, electrostatics, and solvation.

$$E_{total} = E_{bond} + E_{angle} + E_{torsion} + E_{vdw} + E_{coulombic} \quad (1)$$

The interaction energy was calculated as the difference between interaction energies of the complex and of uncomplexed TrxR and Trx, using AMBER force fields.

Interaction energy = energy of complex – energy of uncomplexed TrxR –
energy of uncomplexed Trx (2)

DR TrxR interface residues that were deemed as important for interaction (M84A, K131A, F148A, and F137A), were mutated to alanine using the Biopolymer module of SYBL8.0. The resulting mutants were then minimized and interaction energy was calculated as described for wild-type complex. The difference in interaction energy (ΔInteraction energy) between the wild-type complex and each mutant complex was then calculated as:

$$\Delta\text{Interaction energy} = \text{Interaction energy}_{(\text{mutant})} - \text{Interaction energy}_{(\text{wild type})} \quad (3)$$

CHAPTER THREE: RESULTS AND DISCUSSION

3.1 PROTEIN OVER-EXPRESSION, PURIFICATION AND CHARACTERIZATION

3.1.1 Protein Over-Expression and Purification

Trxs from *E. coli* and *D. radiodurans*, and TrxR from *D. radiodurans* were cloned into a modified pET-30b vector, which can over-express proteins fused to an N-terminal hexa-histidine tag. All three proteins were thus purified by immobilized metal ion affinity chromatography (IMAC) and found to be highly pure after single step purification, as judged by SDS-PAGE analysis (Figure 3.1). EC TrxR and HP TrxR in the ampicillin-resistant vector PROK-1 were over-expressed and purified using ion exchange and affinity chromatography, as described in the Materials and Methods section.

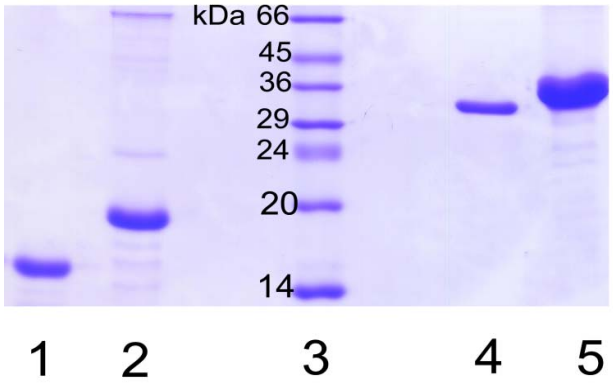


Figure 3.1: SDS-PAGE of purified samples of various TrxR and Trx proteins. (1) EC Trx (2) DR Trx1 (3) LMW marker (4) EC TrxR and (5) DR TrxR.

3.1.2 Estimation of Protein Concentration

After purification, the proteins were dialyzed and concentrated as described in the Materials and Methods section. The concentration of the pure protein samples was determined by the Bradford assay (Bradford, 1976) using bovine serum albumin as a standard (Table 2.2). All the proteins typically yielded 10 mg of pure protein from a 1 L culture. The proteins were subsequently aliquoted into PCR tubes (100μL), flash-frozen in liquid nitrogen, and stored at – 80^{o}C.

3.2 ENZYME KINETICS

3.2.1 Activity Dependence on Cofactor Concentration

Various NADPH concentrations have previously been used (10-500 μM) to perform NADPH reduction and insulin reduction assays (Holmgren & Morgan, 1976; Luthman & Holmgren, 1982). Therefore in order to find the optimum amount of NADPH needed for our assays (insulin reduction and DTNB reduction assays described in section 2.6), initial velocities were measured with various concentrations of NADPH (10-800 μM). The DTNB assay showed saturation at 100 μM of NADPH, but there was an unexpected drop of initial velocity with the insulin reduction assay above 200 μM of NADPH. No activity could be observed with the insulin reduction assay above 600 μM NADPH (Figure 3.2).

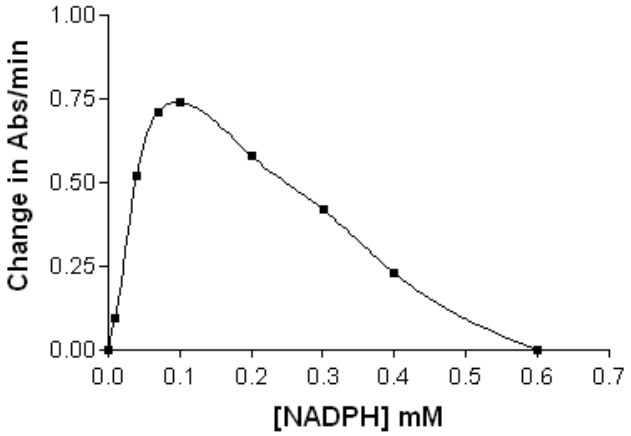
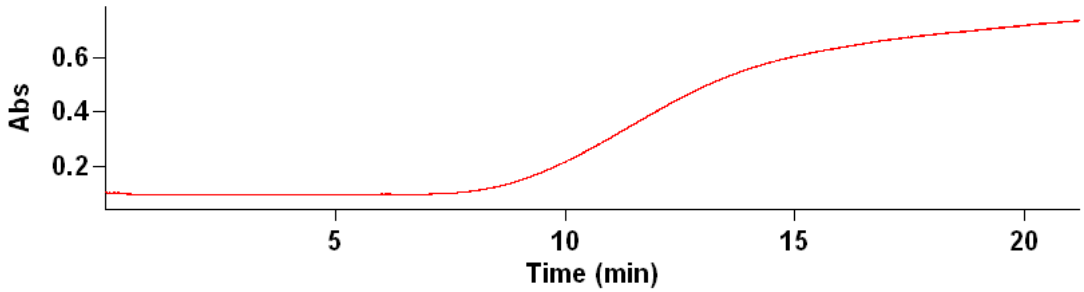
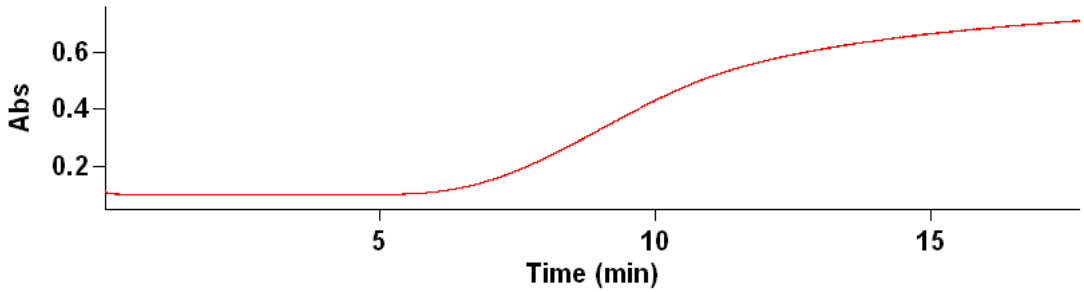


Figure 3.2: A plot of initial velocities vs. cofactor concentration for insulin reduction assay. Initial velocities were measured by monitoring co-factor oxidation at 340 nm under assay conditions described in Materials and Methods section. When increasing amounts of NADPH were used, there was an increase in activity up to a concentration of 0.1mM adding NADPH above 0.1mM results in a decrease in activity.

This does not appear to be a result of substrate inhibition, because substrate inhibition in most cases is partial inhibition i.e. does not approach zero even at high substrate concentrations (Lin et al, 2001). Since this trend was only observed with the insulin reduction assay, the possibility that the observed inhibition was the result of interference of insulin/Trx interaction by NADP/NADPH was tested, using the insulin precipitation assay (Figure 3.3). Carrying out the insulin precipitation assay in the presence of 0.1 mM NADP or 0.1 mM NADPH (0.1 mM was the NADPH concentration used for all the enzymatic assays) does not result in any inhibition of the assay (Figure 3.3). Another possibility is that in the presence of insulin and at high NADPH concentrations, there is alteration of the FO/FR equilibrium of TrxR leading to inhibition. However, it is not clear why this trend would only be observed with the insulin reduction assay. Nevertheless, it was determined that 0.1 mM is the optimum NADPH concentration for both assays.



A)



B)

Figure 3.3: Plot of DTT dependent reduction of bovine insulin; in the absence of NADPH (A) or the presence of NADPH (B). The insulin precipitation assay was carried out by addition of 5µM of Trx1 to a reaction mixture containing 0.85mM insulin and 1mM DTT in 50mM potassium phosphate buffer pH 7.5.

3.2.2 Activity of DR Trx1

All Trxs reported so far can effectively reduce insulin disulfides. Upon reduction, the A and B-chain of insulin dissociate, resulting in aggregation and precipitation of the B-chain. The resulting turbidity can be used to measure the activity of Trx. The standard insulin precipitation assay was performed as previously described (Holmgren, 1979), to confirm the identity of DR Trx1. The addition of 5 µM of DR Trx1 to the reaction mixture described above resulted in the precipitation of insulin after 16 minutes, confirming the identity of the protein. The results

suggest that DR Trx1 reduces insulin in a similar manner to EC Trx1 (7.8 μM of EC Trx1 resulted in rapid precipitation of insulin after 9 minutes) (Holmgren, 1979). Examination of the completed *D. radiodurans* genome revealed two Trxs; annotated as Trx and Trx1 (White et al, 1999). Several other bacteria also contain at least two Trxs usually designated as Trx1 and Trx2. In a classic Trx system, Trx2 has slightly lower redox activity than Trx1 and contains extra N-terminal cysteines (Figure 3.4) which have been suggested to play a role in regulating Trx activity (Miranda-Vizuete et al, 1997).

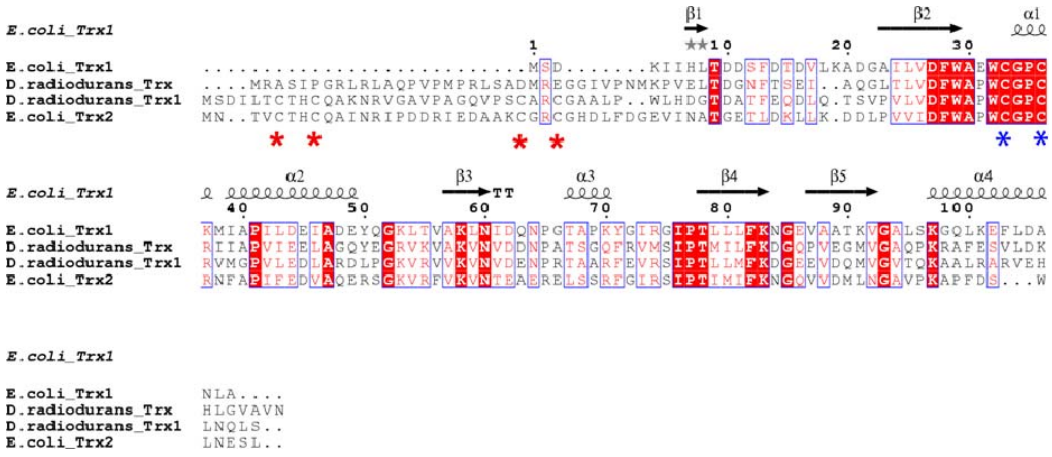


Figure 3.4: Sequence alignment of Trxs from *E. coli* and *D. radiodurans*. The N-terminus of *D. radiodurans* Trx1 is more similar to *E. coli* Trx2 than *E. coli* Trx1. Absolutely conserved active site cysteines are shown in blue (*) whereas N-terminus conserved cysteines of *D. radiodurans* Trx1 and *E. coli* Trx2 are shown in red (*).The figure was produced using the program ClustalW version 2.0 (Larkin et al, 2007) and the ESPript server (Gouet et al, 2003).

The extra N-terminal cysteines (CxxC motifs) have been shown to act as zinc fingers in at least one Trx2. However the role of these zinc binding motifs is not clear, as deletion of these zinc fingers does not affect the oxidoreductase activity of Trx2 (Ye et al, 2007). DR Trx1 contains extra N-terminal cysteines whereas DR Trx does not have any cysteines on the N-terminus; suggesting DR Trx1 is more similar to classic Trx2 than Trx1.

3.2.3 Steady-state Kinetic Analysis of DR TrxR and DR Trx1

The interactions between DR TrxR and Trx1 were analyzed by steady state kinetics using the two enzymatic assays described in the Materials and Methods section. The redox activity of the *E. coli* thioredoxin system was also determined for comparison purposes.

(i) **Insulin Reduction Assay:** This assay examines the ability of thioredoxin to act as a substrate for the reduction of the enzyme thioredoxin reductase. The reaction is monitored by observing the disappearance of NADPH absorbance at 340 nm (Figure 3.5).

(ii) **DTNB Reduction Assay:** This assay is similar to the insulin reduction assay except that DTNB is used as the final recipient of electrons, instead of insulin. The reaction is monitored by measuring the reduction of DTNB at 412 nm (appearance of TNB).

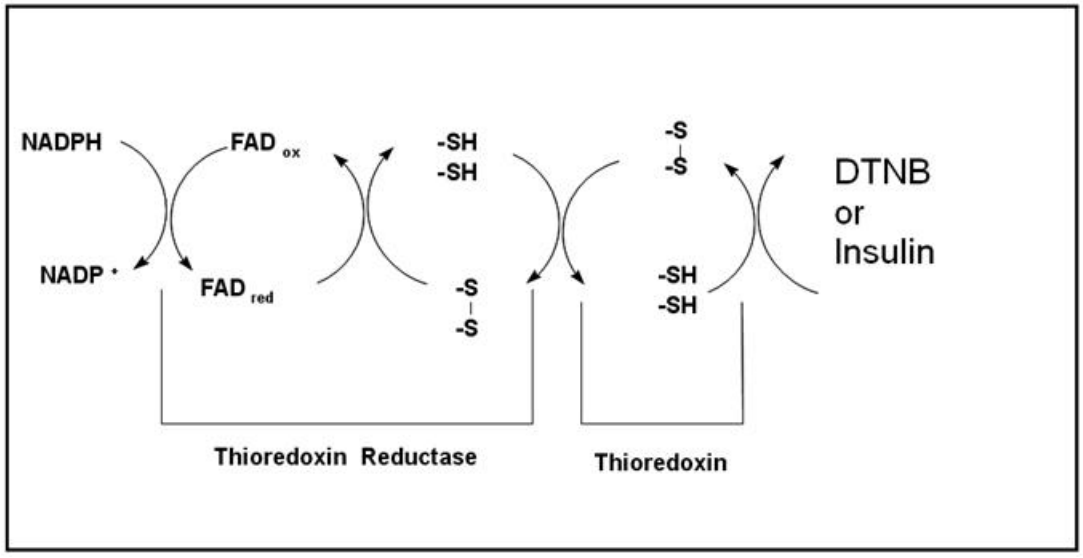


Figure 3.5: Illustration of the flow of reducing equivalents from NADPH to insulin or DTNB. The electron flow is shown for both insulin and DTNB reduction assays.

Both assays are known to have some limitations; DTNB inactivates rat TrxR in the presence of Trx (Luthman & Holmgren, 1982) and HP TrxR reduces insulin even in the absence of Trx

(Pervin, 2006). Fortunately, none of these problems were observed with the *E. coli* or the *D. radiodurans* Trx system. Therefore, both assays were used for kinetic analysis.

A sample saturation curve is shown in Figure 3.6 and the kinetic data are summarized in Table 3.1. The values for K_m and V_{max} were estimated by a non-linear regression fit of the Michaelis-Menten plot where

$$v = \frac{V_{max}[S]}{K_m + [S]} \quad (5)$$

using the Prism software (GraphPad, Software Inc.). There is a slight difference (two fold) between the K_m values of *D. radiodurans* Trx system (5.7μM) and *E. coli* Trx system (2.7μM). There are also slight differences in the overall catalytic efficiency (k_{cat}/ K_m) of the two Trx sytems; the *E. coli* Trx system shows a higher (two to six fold) catalytic efficiency (1.3×10^6 M^{-1} s^{-1}) than the *D. radiodurans* Trx system (1.9×10^5 M^{-1} s^{-1}). These results suggest that abundance rather than efficiency of some ROS scavenging enzymatic systems may be responsible for the higher ROS scavenging ability seen in *D. radiodurans* cell extracts compared to *E. coli* cell extracts (Tian et al, 2007).

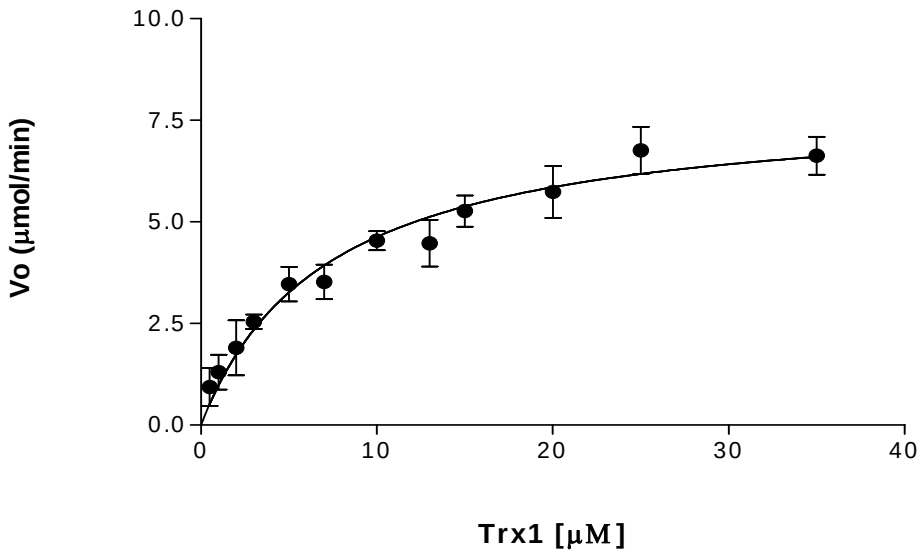


Figure 3.6: Michaelis–Menten plot of DR TrxR using DR trx1 as substrate. The reaction was monitored by consumption of NADPH via a decrease in absorbance at 340 nm. Initial velocity was measured as µmoles of NADPH consumed per minute. Reaction mixtures contained 100 mM HEPES buffer (pH 7.4), 2 mM EDTA, 30 µM solution of bovine insulin (Sigma), 0.1 mM NADPH (Calbiochem), 0.1 µM DR TrxR and DR Trx1 (0-35 µM).

Assay	K_m **(µM)**	k_{cat} **(s^{-1})**	k_{cat}/K_m **(M^{-1} s^{-1})**
DR TrxR/DR Trx1	5.7 ± 1.9 (2.75 ± 0.40)	1.1 ± 0.1 (9.4 ± 1.0)	1.9×10^5 (3.4×10^6)
DR TrxR/EC Trx1	32.4 ± 8.5	18.0 ± 2.6	5.5×10^5
EC TrxR/DR Trx1	44.4 ± 5.5	4.6 ± 0.2	1.0×10^5
EC TrxR/EC Trx1	2.7 ± 0.7 (0.7 ± 0.003)	3.6 ± 0.01 (4.4 ± 0.04)	1.3×10^6 (6.3×10^6)

Table 3.1: Kinetic constants for *D. radiodurans* Trx system. Both DR trx1 and EC Trx1 are used as substrates. The k_{cat} values for the *E. coli* system clones are about six fold lower than the literature values; using both the insulin reduction and DTNB reduction assays (in brackets) (Holmgren & Morgan, 1976; Luthman & Holmgren, 1982).

3.2.4 Cofactor Specificity

It has been reported that DR TrxR may have dual co-factor specificity, i.e. it can utilize NADPH or NADH for its redox activity. In that study, the DTNB reduction assay was performed on cell free extracts from *D. radiodurans*, without purification of either TrxR or Trx (Seo & Lee, 2006). To test the possibility that DR TrxR has dual cofactor specificity, the insulin reduction assay was performed using both NADPH and NADH as a source of reducing equivalents as described in the Materials and Methods section. We determined that DR TrxR can only utilize NADPH as a source of reducing equivalents. No activity was seen when NADH was used as a reductant for the insulin reduction assay (some trace activity was seen with the DTNB reduction assay) (Figure 3.8). Sequence analysis of DR TrxR also suggests NADPH-dependence (Figure 3.7). DR TrxR has the GXGXXA motif (common in NADPH-dependent enzymes) instead of the GXGXXG motif (common in NADH-dependent enzymes). Our results are consistent with previous studies which showed that bacterial TrxRs had no activity (Baker et al, 2001a) or very low activity (Thelander, 1967) when NADH is used. Therefore, it can be concluded that the DTNB reduction observed from *D. radiodurans* cell free extracts is probably due to a source other than the *D. radiodurans* TrxR.

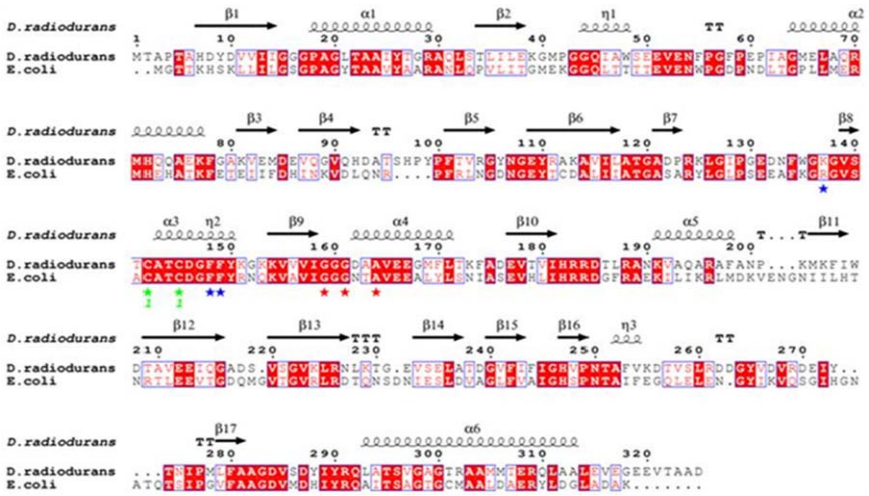
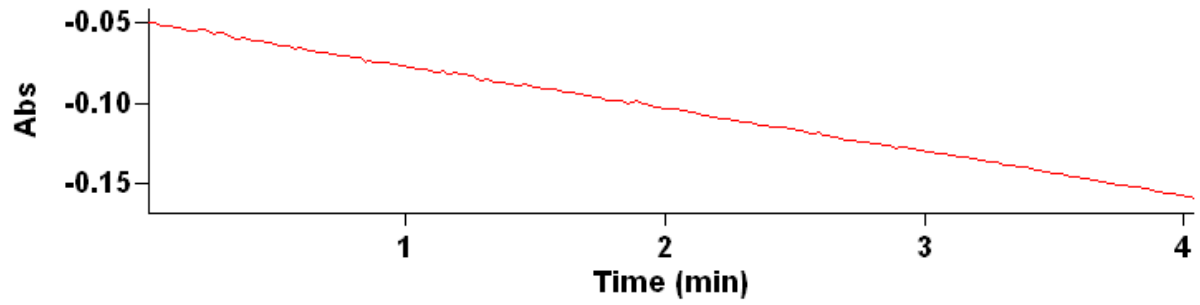
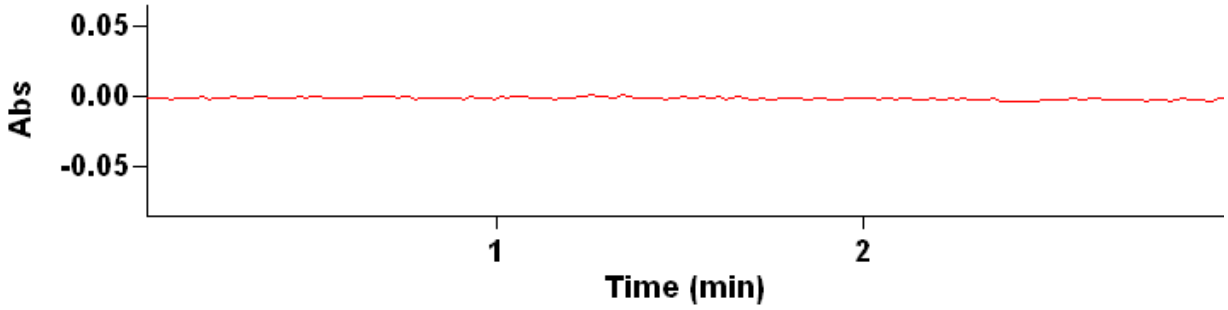


Figure 3.7: Sequence alignment of DR TrxR and EC TrxR. The conserved active-site cysteine residues are indicated with green asterisks, the key residues involved in interaction with Trx are indicated with blue asterisks, and the GXGXXA motifs are indicated with red asterisks. The figure was produced using the program ClustalW version 2.0 (Larkin et al, 2007) and the ESPript server (Gouet et al, 2003).



A)



B)

Figure 3.8: Insulin reduction assay using NADPH or NADH as the cofactor. A) NADPH used as cofactor B) NADH used as cofactor. No oxidation of cofactor was observed when NADH was used as a reductant.

3.3 CRYSTALLIZATION

3.3.1 Crystallization of DR TrxR

Crystals appeared in several conditions after 3 days using sitting drop crystallization technique. After optimization, the best crystals were obtained in 35% (w/v) polyethylene glycol (PEG) 4000, 0.2 M ammonium acetate, 0.1 M citric acid buffer, pH 5.1, with dimensions of 0.3 x 0.3 x 0.5 mm (Figure 3.9) after 2 weeks. These crystals diffracted to 2.4 Å, at the in-house X-ray diffraction facility at the Saskatchewan Structural Sciences Center (SSSC) (DX8 Proteum diffractometer). They were deemed highly diffracting crystals; therefore, no further optimization was attempted. In addition, no cryoprotectant was needed during flash freezing with liquid nitrogen, because no water rings were observed in the diffraction pattern (Figure 3.11). The final high resolution diffraction was carried out at beamline 14-BM-C of the APS, at Argonne National Laboratory (Argonne, IL, USA).

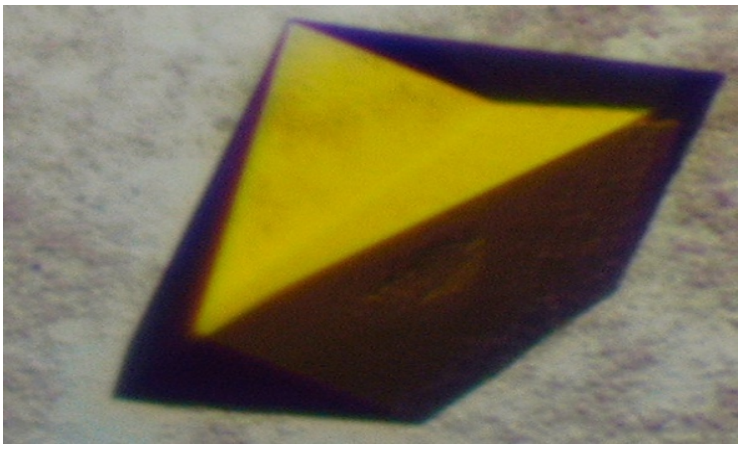


Figure 3.9: The picture of DR TrxR crystal with dimensions of 0.3 x 0.3 x 0.5 mm (Obiero et al, 2006).

3.3.2 Crystallization of HP TrxR

Initial HP TrxR crystals were also obtained from commercial sparse-matrix broad screens (Qiagen), under conditions described in the Materials and Methods section. These crystals initially diffracted to 3.0 Å at the Canadian Light Source. HP TrxR crystals underwent more extensive optimization compared to DR TrxR crystals. Optimization for HP TrxR included varying precipitant concentrations (ammonium sulfate), additives, temperatures, and crystallization methods (vapor diffusion and microbatch). After optimization, large crystals (0.2 mm x 0.2 mm x 0.3 mm) with sharp edges (Figure 3.10) were obtained in one week at 4 ºC with the hanging drop crystallization technique. The best crystallization condition contained 0.2M ammonium nitrate, 0.2M spermine, and 2.0M ammonium sulfate. These crystals diffracted to 2.4 Å at the Canadian Light Source.

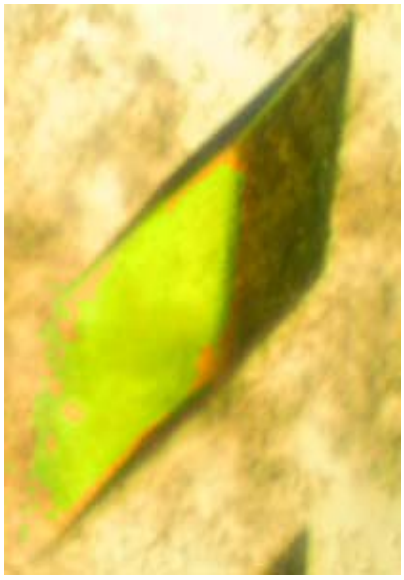


Figure 3.10: The picture of HP TrxR crystals. The crystals were grown by hanging drop vapor diffusion crystallization method using 0.2 M ammonium nitrate, 0.2 M spermine, and 2.0 M ammonium sulfate at 4ºC.

3.4 DATA COLLECTION AND PROCESSING

Data collection and processing for HP TrxR and DR TrxR was performed as described in the Materials and Methods section. Using the R_{merge} values and the self rotation function from MOLREP (Vagin & Teplyakov, 1997), the space group for both DR TrxR and DR TrxR crystals was determined to be $P3_221$. Matthews coefficient calculations (Matthews, 1968) suggested the presence of a dimer in the asymmetric unit of the DR TrxR crystals (with 47.0% solvent content) and a trimer in the asymmetric unit of HP TrxR crystals (with 60.0% solvent content). A summary of crystallographic data statistics is shown in Table 3.2.

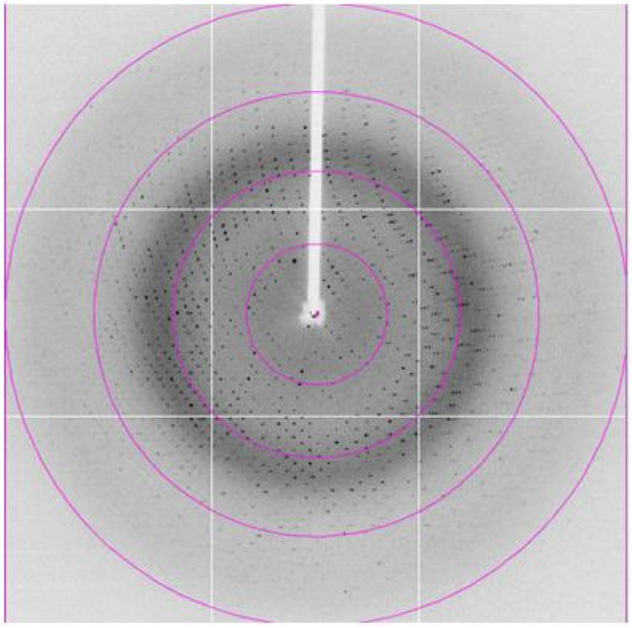


Figure 3.11: A 0.5º oscillation image collected at the APS at 100 K temperature from a *D. radiodurans* TrxR crystal.

	HP TrxR	DR TrxR
	CLS	APS
Data collection		
Temperature (K)	100K	100K
Wavelength (Å)	1.4	0.98
Space Group	$P3_221$	$P3_221$
Unit Cell parameters (Å)	a= 89.3, b= 89.3, c= 280.3 α=β=90, γ=120	a= 84.3, b= 84.3, c= 158.9 α=β=90, γ=120
Resolution limits (Å)	39.56-2.4 (2.66-2.43)	80.06–1.90 (2.00–1.90)
Total number of reflections	524256 (116967)	311938 (43600)
Number of unique reflections	49803 (11595)	49077 (6865)
Completeness (%)	100 (99.6)	93.6 (90.6)
R_{merge} (%)	4.9 (17.0)	6.9 (33.2)
$I/\sigma I$	24.3 (7.2)	21.6 (4.5)
Redundancy	10.6	6.4
Matthews coefficient ($Å^3$ Da^{-1})	3.1	2.3
Refinement		
R_{factor} (%)	18.0	19.2
R_{free}	25.0	24.2
Solvent content (%)	60.6	47.2
Number of molecules in asymmetric unit	3	2
Number of atoms		
Protein	7072	4728
Ligand (FAD)	159	106
Number of water molecules	484	521
Overall B values	39.0	22.0
Average B values		
Main chain	37.2	19.6
Side chains	40.1	23.0
Waters	45.9	40.4
r.m.s deviations		
Bond lengths (Å)	0.026	0.016
Bond angles (°)	2.3	1.7
Ramachandran plot, non-glycine residue in		
Most favourable region (%)	86.7	89.5
Additionally allowed region (%)	12.3	10.2
Generously allowed region (%)	1.0	0.20
Disallowed region (%)	0.0	0.20

*Values in parentheses are for highest resolution shell.

Table 3.2: Data collection and refinement statistics for DR TrxR and HP TrxR

3.5 STRUCTURE DETERMINATION

Molecular replacement was used for structure determination of both DR TrxR and HP TrxR. Two programs were used for molecular replacement, Phaser (McCoy et al, 2007) and Molrep (Vagin & Teplyyakov, 1997) both which are now part of an automated scheme for molecular replacement by use of multiple proteins (MrBUMP) (Keegan & Winn, 2008). Both programs arrived at the correct solution. The rotation function of both Phaser and Molrep generates a large number of different orientations (see Section 1.4.6). These orientations are described by Euler angles. Eulerian angles are usually described by operational order: angle α around the *z* axis, an angle β around a new *y* axis, and finally an angle γ around the new *z* axis. The two programs utilize different approaches for the rotation and translation functions. Molrep uses traditional rotation and translation functions, first proposed by Crowther and Blow (Crowther and Blow, 1967), with some additional features (like the packing function) that provide additional discrimination between correct and incorrect solutions. Phaser on the other hand, uses maximum likelihood techniques for rotation and translation functions. For moderately or highly homologous structures where the probability distributions of molecular replacement solutions are Gaussian (as was the case with DR TrxR and HP TrxR structures) Phaser and Molrep will give similar results. In cases where there is poor homology between known and unknown structures, other probability distributions may exist. In these cases, Phaser will yield superior results to Molrep (because it weights models based on correctness of the model as judged from parameters derived from known protein structures). The correct solution is identified by Z-scores and log likelihood gains (LLG) in Phaser (Drenth, 1999; McCoy et al, 2007).

3.5.1 DR TrxR Structure Determination

The DR TrxR structure was determined using chain B of the *M. tuberculosis* TrxR structure (Protein Data Bank (PDB) code 2A87) as the search model for molecular replacement (Obiero et al, 2006). The correctness of the resulting solution was confirmed from the electron density of FAD (which appeared in the expected positions although it was omitted from the initial search model) (Figure 3.12).

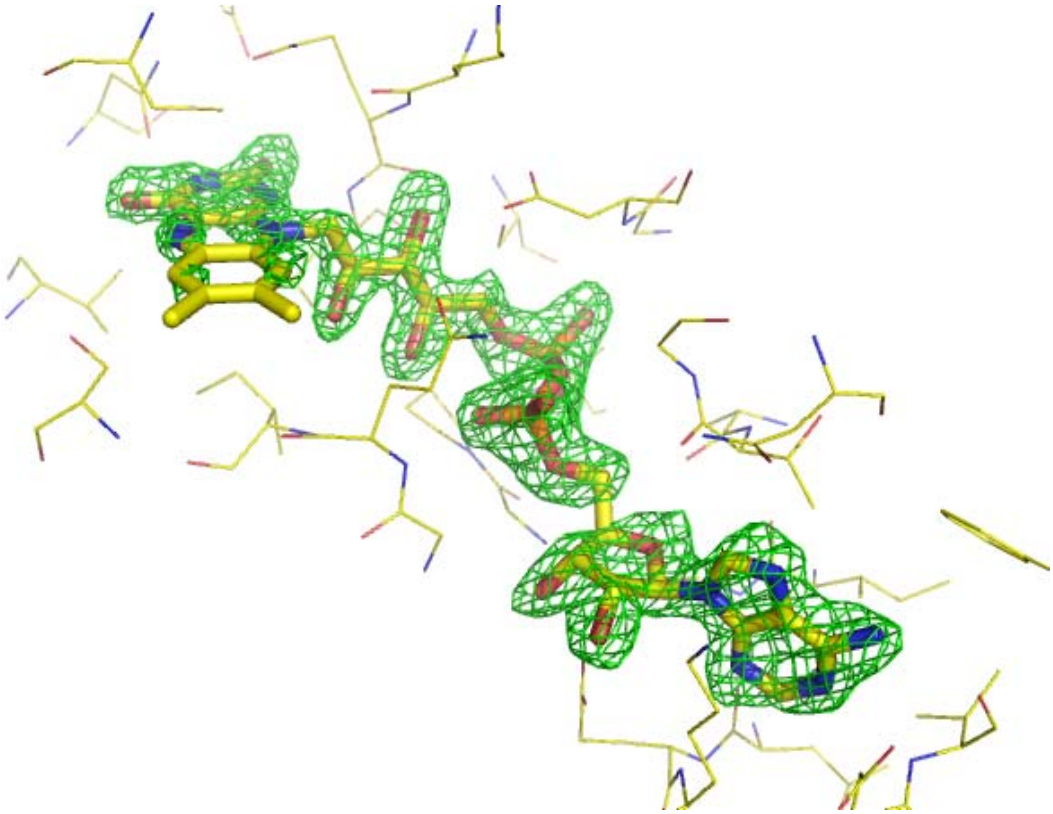


Figure 3.12: Electron density map of the initial solution generated by MrBUMP. Difference map (Fo - Fc) generated from the molecular-replacement solution from MrBUMP. The difference map (coloured green) is contoured at 2.8σ. The actual model is shown as yellow lines, while the FAD (not included in initial model, but inserted to show location) is shown as a stick model. The figure was generated using PyMOL (DeLano, 2002a).

Initial restrained refinement carried out using REFMAC5 (Murshudov et al, 1997) as part of MrBUMP (Keegan & Winn, 2008) resulted in an R_{free} of 0.431, from an initial R_{free} of 0.547. Further refinement was performed using REFMAC5 with manual rebuilding using Coot (Emsley

& Cowtan, 2004). Tight non-crystallographic symmetry restraints for the main-chain atoms and loose restraints for the side chain atoms were applied in the late stages of refinement. FAD was easily modeled into positive difference electron density at the expected sites. The program ARP/wARP (Perrakis et al, 1997) was used to add water molecules and the resulting model manually inspected for correctness using Coot.

3.5.2 HP TrxR Structure Determination

The HP TrxR structure was determined as described for the DR TrxR structure, except that the EC TrxR (PDB code 1TDE) was used as the search model. The Fo-Fc map of the initial solution also showed positive density for FAD, although it was omitted in the initial search model. Initial restrained refinement carried out using REFMAC5 as part of MrBUMP resulted in an R_{free} of 0.471 from 0.537. Further refinement and model building was carried out as described for the DR TrxR structure (Obiero et al, 2006) (Section 3.5.1.).

3.6 MODEL BUILDING, REFINEMENT AND VALIDATION

All refinement programs use constraints or restraints (bond lengths, bond angles, dihedral angles, and planarity) based on small molecule crystal structures or high resolution protein structures. The use of small molecule parameters is based on an assumption that proteins will have the same parameters as small molecule crystal structures. In constrained (rigid body) refinement, only overall translation and rotation can be varied (by varying dihedral angles). This reduces the number of parameters to be refined, but makes it easier to move large parts of the structure. In contrast, restrained refinement allows other parameters such as bond lengths, bond angles, and van der Waals contacts to vary (around a standard value), in addition to the dihedral angles. This makes it easy to move small parts of the structure but difficult to move large parts of the structure.

The aim of all refinement procedures is to obtain a refined structure where the deviation from ideal values is within an acceptable range. Refinement procedures (for the initial models from MrBUMP) were done using a maximum likelihood refinement program, REFMAC5, first with rigid body refinement (20 cycles) followed by several rounds of restrained refinement. The restrained refinement process was interspersed with manual rebuilding using COOT. Tight non-crystallographic symmetry restraints for the main-chain atoms and loose restraints for the side chain atoms were applied in the late stages of refinement. FAD was easily modeled into positive difference electron density at the expected sites using F_o-F_c electron density map. The program ARP/wARP was used to add water molecules and the resulting model manually inspected for correctness using COOT. The final model was examined for the main chain torsion angles (φ, ψ) by the program PROCHECK (Laskowski et al, 1993). A Ramachandran plot produced by PROCHECK is shown in Figures 3.14 and 3.16. Except for DR TrxR's glutamate 271 and the glycine residues, all of the amino acid residues are in the allowed or generously allowed regions. The refinement and final model statistics for the final models are shown in Table 3.2.

3.7 THIOREDOXIN REDUCTASE STRUCTURES

3.7.1 Oxidized DR TrxR Structure

The structure of oxidized DR TrxR was refined to a 1.9Å resolution. The final restrained refinement resulted in an overall crystallographic *R*-factor of 19% and an R_{free} of 24%. All the residues of the polypeptide chain of both subunits (A and B) were well defined in the electron density maps except for terminal residues of both the C-terminal and the N-terminus. One residue from the N-terminus and 11 residues from the C-terminus were not observed in electron density for subunit A and five residues from the N-terminus and 11 residues from the C-terminus

were not observed in the electron density for subunit B. The FAD molecules were also well defined in the electron density for both subunits (A and B).

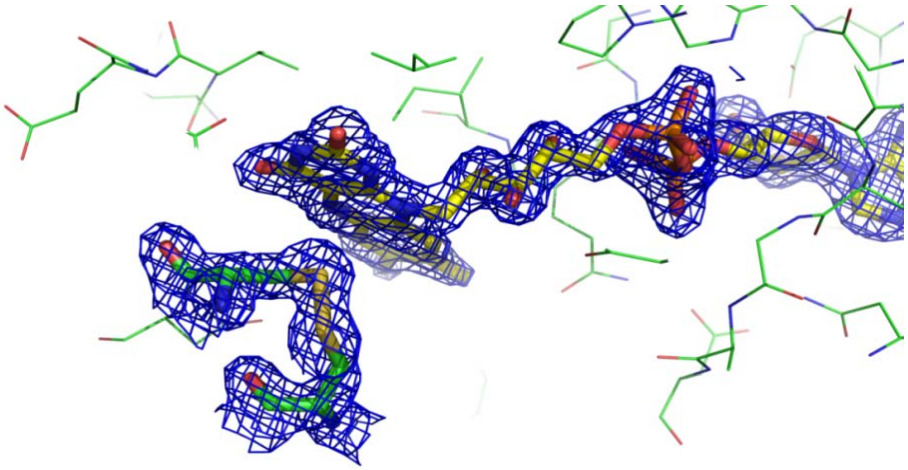


Figure 3.13: Electron density map (2fo-Fc map) of the active site disulfide in the FO conformation. In the FO confirmation, the active site disulfide is close to the isoalloxazine ring of FAD. Under oxidizing conditions the isoalloxazine ring of FAD is flat.

The active site residues are located on a loop that connects a β-strand and an α-helix (loop β8–α3 contains the first active-site cysteine residue C142, and helix α3 contains the second cysteine residue Cys145) (Figure 3.18). The electron density map of the active site disulfide (C142, C145) is shown in Figure 3.13. The stereochemistry of the model is good as judged by the Ramachandran plot (Figure 3.14). About 90% of the residues lie within the most favored region of the Ramachandran plot, with the rest of the residues in the additionally allowed regions of the Ramachandran plot. One residue (E271) is well defined in the electron density maps, but it is in the disallowed region of the Ramachandran plot. This residue is part of a tight turn (γ-turn) involving three residues D270, E271, and I272. These three residues are not conserved (Figure

3.18) across bacterial TrxRs and there is no apparent biological role for the tight turn. The residue (E271) participates in dimer formation just like the equivalent residue from HP TrxR (F267) which also participates in dimer formation although it is part of a loop (Figure 3.20). The asymmetric unit contained two monomers forming a homodimer as found in solution. The final results of the refinement are given in Table 3.2.

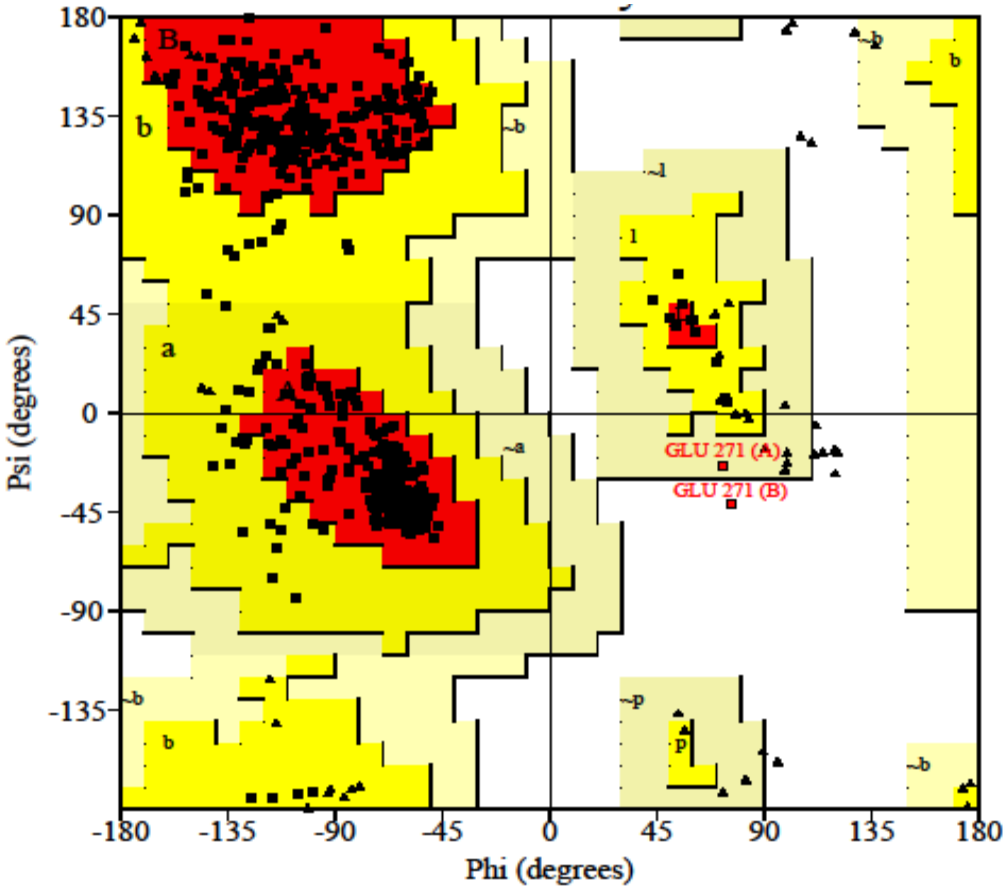


Figure 3.14: The Ramachandran plot of DR TrxR. Residues in most favored regions (red), additional allowed regions (deep yellow), and generously allowed regions (light yellow) are 89.7%, 10.0%, and 0.2% respectively All the residues lie within the most favored or additionally allowed regions except glutamate 271 which is a tight turn involving three residues D270, E271, and I272 (Obiero et al, 2010).

Overall the two monomers are very similar (they superpose with a root mean square deviation (rmsd) of 0.45 Å), but several loop regions show slightly different conformations, with the largest shift of 2.5 Å found at residue 229. Each monomer is composed of two domains, the FAD-binding domain (residues 1 to 123 and 249 to 325) and NADPH-binding domain (residues 124 to 248). Both domains consist of variants of the canonical nucleotide binding fold that was first seen in glutathione reductase (Schulz et al, 1978). A typical Rossman fold consists of a slightly twisted parallel β-sheet surrounded by α-helices on both sides. However, the FAD-binding and NADPH-binding domains of DR TrxR contain both parallel and anti-parallel β-sheets, surrounded by α-helices on one side only. Each domain contains two β-sheets: a central five-stranded parallel β-sheet and a three-stranded anti-parallel β-sheet, packed against each other. The two β-sheets are surrounded by several α-helices on one side (Figure 3.15).

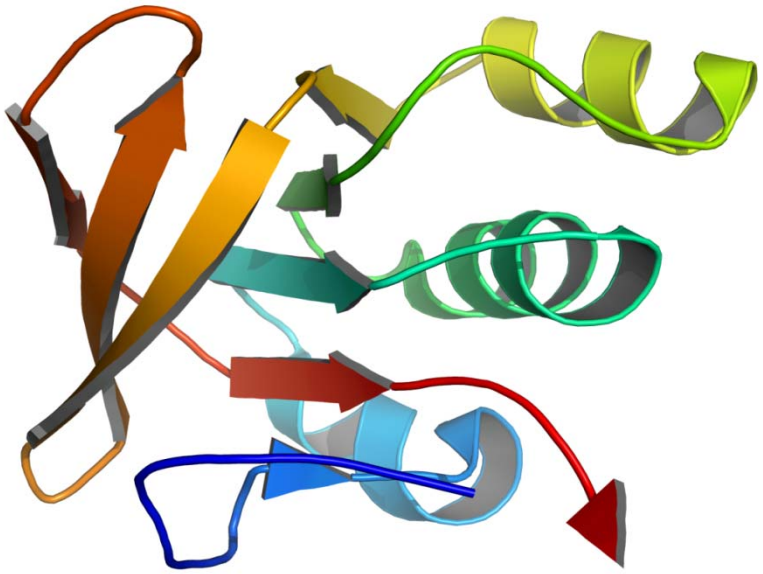


Figure 3.15: The NADP(H)-binding domain of DR TrxR. Two β-sheets are surrounded by α-helices on one side.

3.7.2 Oxidized HP TrxR Structure

The structure of oxidized HP TrxR was refined to a 2.4Å resolution. The final restrained refinement resulted in an overall crystallographic *R*-factor of 20% and an R_{free} of 26%. Unlike the DR TrxR where there are missing terminal residues, all 311 residues of HP TrxR are well defined in the electron density in all the three subunits (the asymmetric unit contained three monomers A, B, & C). Much like the DR TrxR, the FAD molecules were also well defined in the electron density for all chains. The stereochemistry of the HP TrxR model was also good, as judged by the Ramachandran plot (Figure 3.16).

About 87% of HP TrxR residues lie within the most favored region of the Ramachandran plot, with the rest of the residues in the additionally allowed regions. The HP TrxR structure lacks the tight turn seen in DR TrxR; it has a loop instead. The overall structure is however very similar to the DR TrxR structure, with all subunits containing both the FAD and NADPH domain. Each domain in turn is consisted of a variant of the Rossman fold described in section 3.7.1. The final results of the refinement are given in (Table 3.2). Overall the three monomers are very similar (they superpose with an rmsd of 0.45 Å), but several loop regions show slightly different conformations.

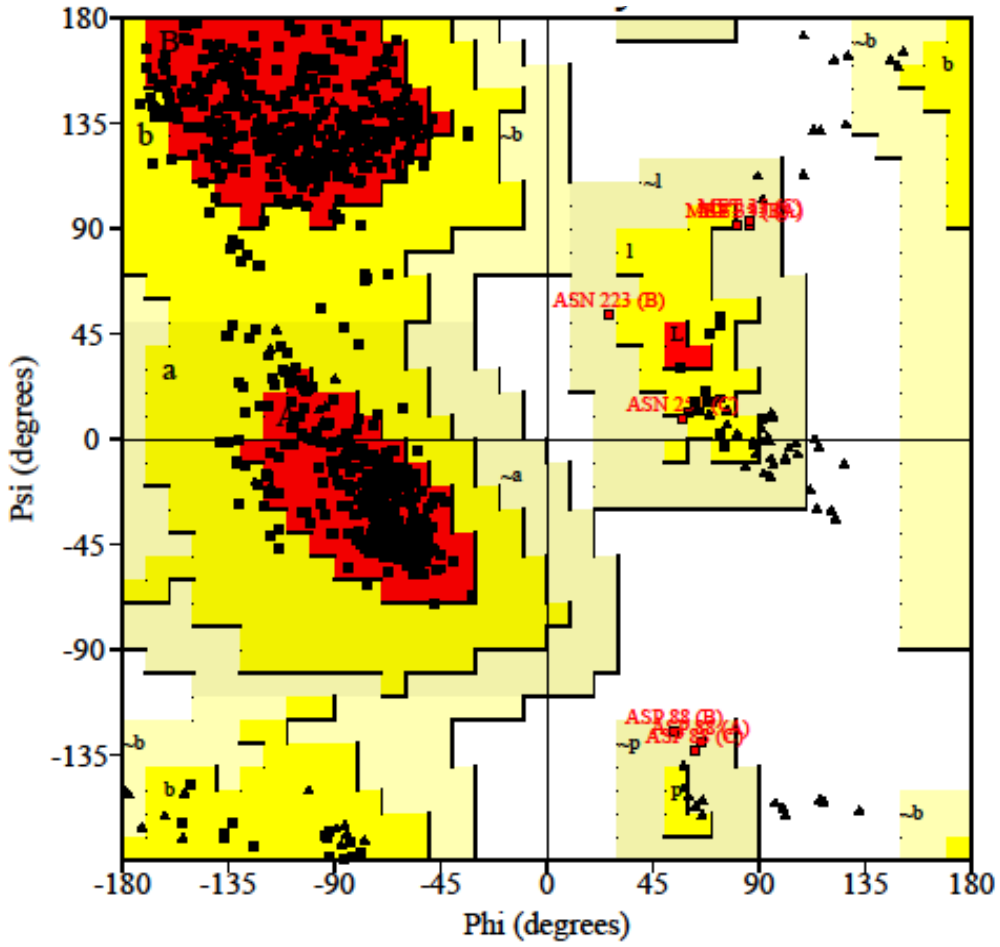


Figure 3.16: The Ramachandran plot of HP TrxR. All the residues lie within the most favored or generously allowed. Residues in most favored regions (red), additional allowed regions (deep yellow), and generously allowed regions (light yellow) are 86.7%, 12.3%, and 1.0% respectively. No residues lie in the disallowed regions.

3.7.3 FAD Binding

The cofactor FAD is tightly but non-covalently bound to both DR TrxR and HP TrxR. All bacterial TrxRs characterized to date also contain tightly bound FAD. This suggests that FAD is incorporated during protein folding and may play a role in protein folding. A protein containing oxidized FAD can be easily identified by the bright yellow color of the protein solution or crystals. The FAD binding modes of DR TrxR and HP TrxR are similar to each other, and to those observed in other enzymes belonging to the pyridine nucleotide–disulfide oxidoreductase family. FAD binding to TrxR can be divided into three regions: the riboflavin binding region,

diphosphate binding region, and adenosine binding region. The FAD binding domain of bacterial TrxRs is divided into two parts (section 3.7.1). Most of the interactions with FAD occur within the N-terminal part of the FAD binding domain. The C-terminal part of the FAD domain only interacts with the isoalloxazine ring of the riboflavin portion of FAD.

i) Riboflavin binding region

The riboflavin binding region consists of two parts the isoalloxazine ring and a ribityl moiety. The isoalloxazine ring and the ribityl moiety are completely buried in the protein and are dominated by hydrogen bonding interactions with the main chain and side chains of two residues, N54 and L293 (the amino acid numbering scheme is based on DR TrxR for the discussion in this section; corresponding HP TrxR amino acids are shown in Figure 3.18). N54 is invariant and L293 has conserved substitutions. In addition, several water molecules participate in a hydrogen-bond network.

ii) Diphosphate binding region

The oxygen atoms of the diphosphate group interact primarily with main-chain nitrogen atoms, except for the hydrogen bond to the side chain of Q45 (Figure 3.18). It is worth noting that there is no positively charged residue in the immediate vicinity of the diphosphate group. Instead, the negative charge on the phosphate groups is stabilized by the dipole of helix α1 and hydrogen bonding to the main-chain nitrogen atoms. Water molecules also participate in the hydrogen-bond network.

iii) Adenosine binding region

The adenosine binding region is composed of two parts, an adenine and a ribose sugar binding regions. The adenine ring binds in a pocket formed by residues G15, E86, V87, N250, and F253; but does not form any hydrogen bonding with any main chain or side chain atoms except

between the main chain oxygen of V87 and the amine of the adenine ring (Figure 3.17). Both hydroxyl oxygen atoms of the adenosine ribose form hydrogen bonds with the carboxyl group of E38. Although residue E38 is not conserved (Figure 3.18), it has been observed that all bacterial TrxRs form H-bonds with both hydroxyl oxygen atoms of the adenosine ribose (Figure 3.17). In EC TrxR for example, hydrogen bonding to both hydroxyl groups is formed by side chains of S14 and T36. This feature has been observed in many nucleotide-binding proteins (Karplus & Schulz, 1989; Mattevi et al., 1992), but not in mammalian TrxR, where there is only one hydrogen bond formed with the adenosine ribose.

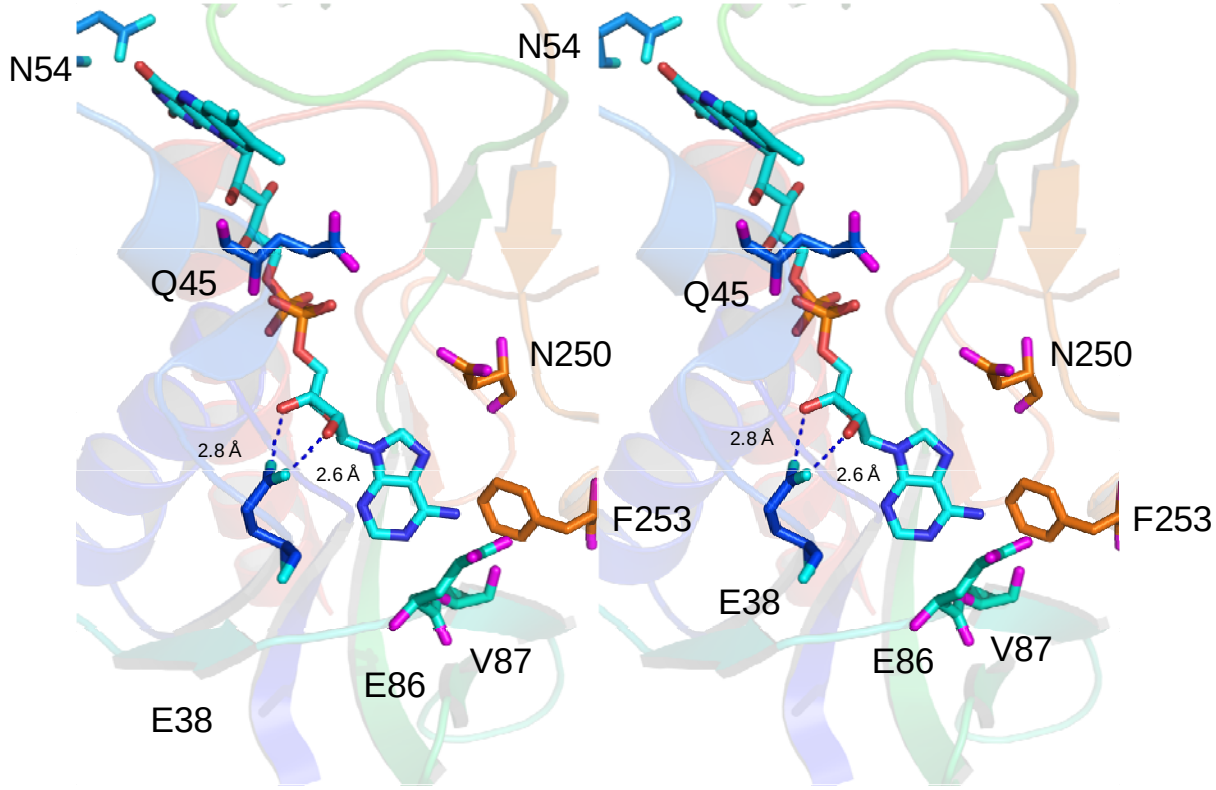


Figure 3.17: Stereoview of key residues that interact with FAD in the DR TrxR structure. Residues N250, F253, V87, and E86 form a binding pocket for adenine ring and E38 forms hydrogen bonds with both hydroxyl oxygen atoms of the adenosine ribose.

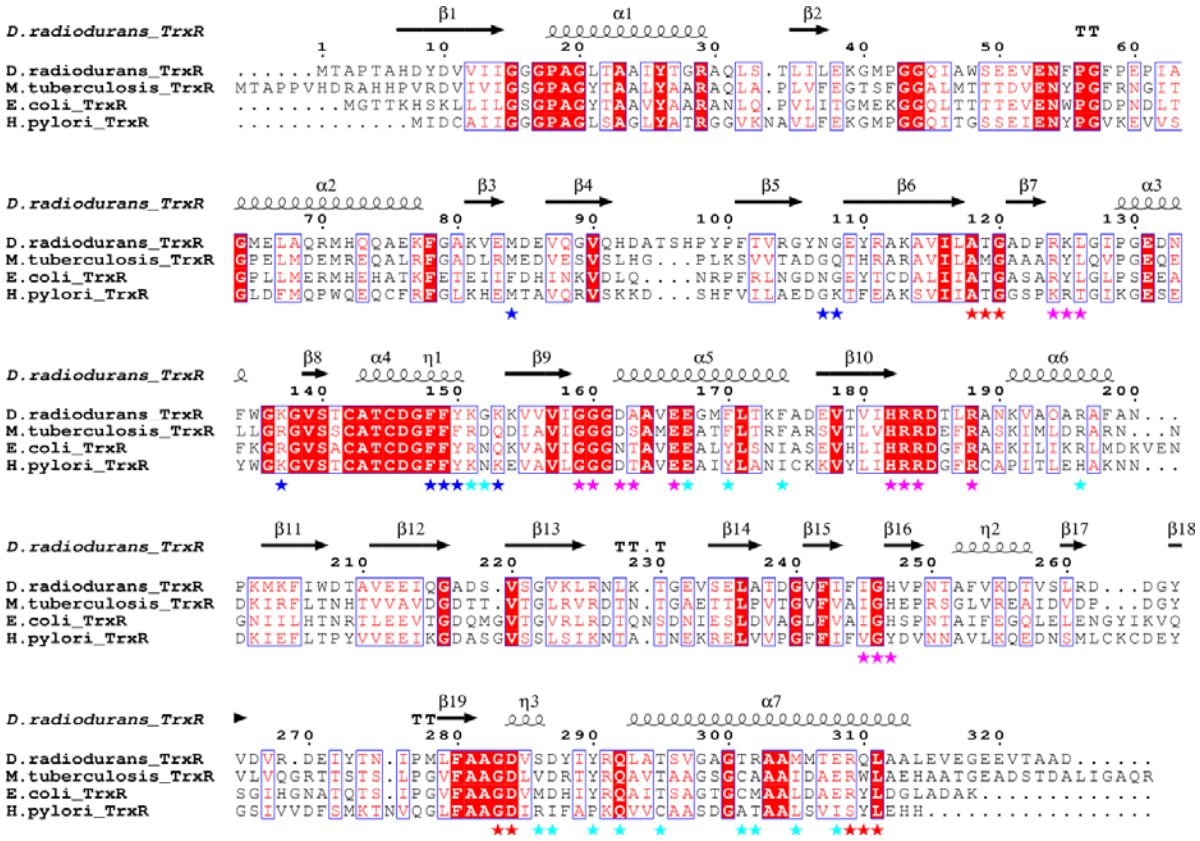


Figure 3.18: Sequence alignment of various TrxRs showing thioredoxin binding, nucleotide binding, and dimer formation residues. Invariant residues are highlighted by a red background. The secondary structure of *D. radiodurans* TrxR assigned by the ESPript server is shown on the top line. Possible functions of individual residues are indicated on the bottom line with asterisks: FAD-binding (red), NAD(P)H-binding (pink), dimer formation (cyan), and Trx-binding (blue). The numbers on the top line represent *D. radiodurans* amino acids. The figure was produced using the program ClustalW and the ESPript server.

3.7.4 NADPH binding

The binding mode of NADPH has been determined in several bacterial TrxR structures; both in the FO conformation and the FR conformation of the enzyme. The binding mode is similar in both conformations, i.e. NADPH-protein interactions, despite the fact that the orientation of the NADPH-binding domain relative to the FAD-binding domain is very different in both conformations. However there are small differences between the NADPH-binding modes of FO and FR conformations. NADPH is more exposed to solvent in the FO conformation than in the FR conformation. The NADPH binding residues are mostly conserved among bacterial TrxRs (Figure 3.18). Among the conserved residues, there are three invariant arginine residues (DR TrxR residues: R183, R184, and R188) that bind the 2'-phosphate of NADP (hydrogen bond/salt bridges). These three residues may be responsible for the ability of bacterial TrxR reductases to discriminate between NADPH and NADH (Gustafsson et al., 2007).

The HP TrxR structure (PDB code: 3ISH) (without NADP) appears to be remarkably similar to other HP TrxR structures (with NADP bound PDB codes: 2Q0L, 2Q0K; Gustafsson et al, 2007) despite the presence of the additional molecules (NADP) in other HP TrxR structures. All the C^{α} atoms of our HP TrxR (3ISH) can be superposed to those of NADP-bound HP TrxR structures (2Q0L, 2Q0K) with an rmsd of ~0.4 Å. NADPH-binding does not alter the main chain, however, the side chains involved in interaction with 2'-phosphate of NADPH have different orientations compared to those of NADP-bound HP TrxRs (Figure 3.19). Overall the HP TrxR NADPH domain has higher temperature factors (43 $Å^2$) compared to the FAD domain (35 $Å^2$). The disparity of temperature factors in bacterial TrxRs is thought to be due the flexibility of NADPH domain, which is in equilibrium between the FO and FR conformations in solution (Akif et al, 2005).

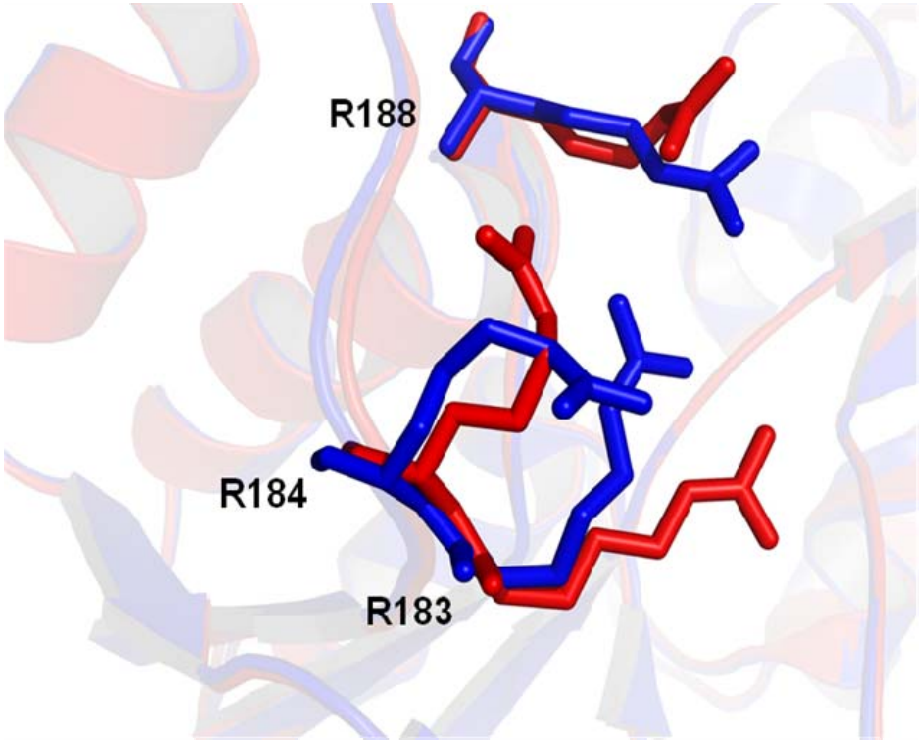
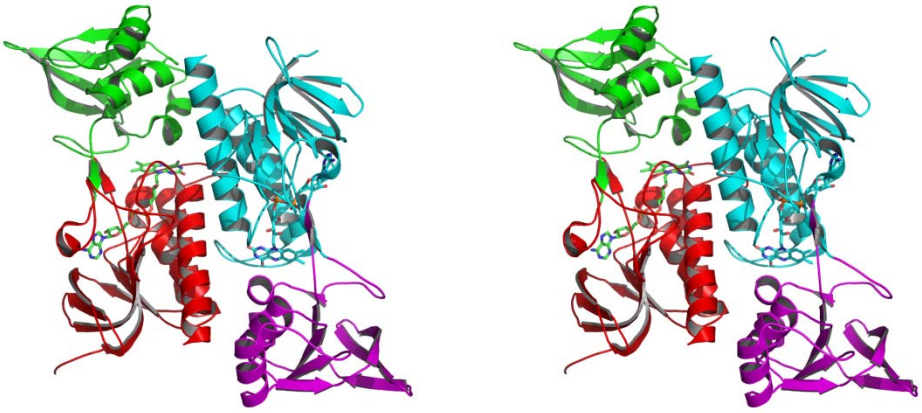


Figure 3.19: HP TrxR structures showing NADPH diphosphate binding residues. HP TrxR structures are superposed and shown with NADPH (Red) and without (Blue) NADPH.

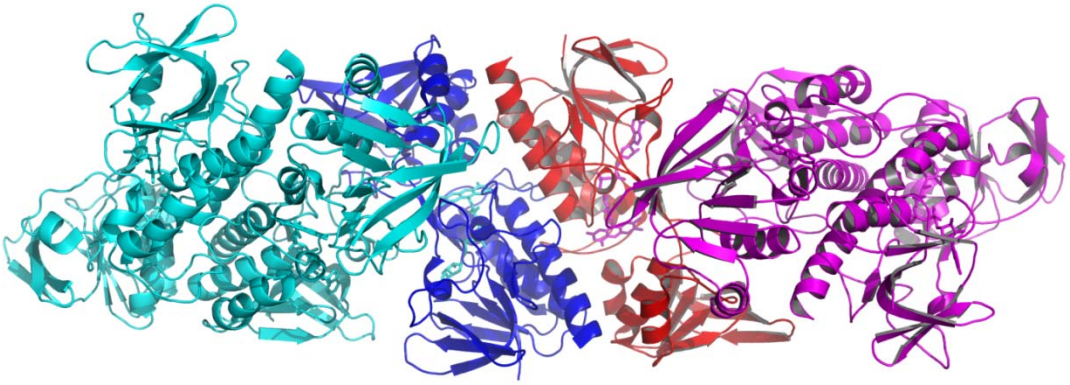
3.7.5 Dimer Interface

Bacterial TrxRs are homodimers in solution. The asymmetric unit of DR TrxR contained a dimer whereas the asymmetric unit of HP TrxR contained a trimer (consisting of the correct biological dimer and a monomer which generates a dimer across the two-fold of the $P3_221$ space group given the correct transformation matrix) (Figure 3.20). Unlike eukaryotic TrxR, which has an interface domain that facilitates dimerization, bacterial TrxR utilizes residues from both NADPH binding domain and FAD binding domain for dimerization. Residues that participate in the dimerization of bacterial TrxR were identified (Figure 3.18) by visual inspection of the interfaces and by analysis using various servers (Proface and Robetta) (http://www.boseinst.ernet.in/resources/bioinfo/stag.html; http://robetta.bakerlab.org/alascansubmit.jsp).

Most of these residues are conserved or contain conservative substitutions across bacterial TrxRs (Figure 3.18). The interface contains almost an equal number of polar (including charged residues) and non-polar residues. However, the Robetta server predicts that non-polar residues contribute more to interface stability than hydrophilic residues, as evidenced by higher $\Delta\Delta G$ values upon mutation to alanine. This prediction is consistent with PDB surveys (of other permanent complexes) which suggest that hydrophobic effects play the dominant role in interface stability of most protein complexes (Jones & Thornton, 1996; Moreira et al, 2007; Young et al, 1994).

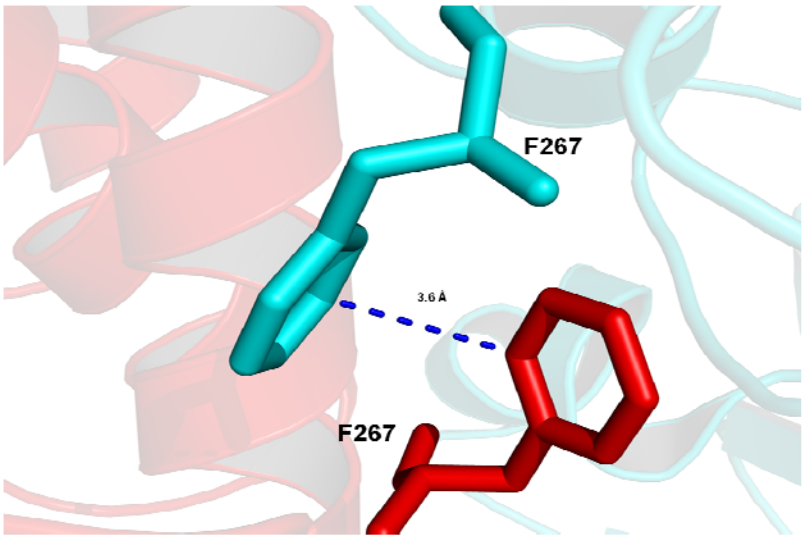


A)

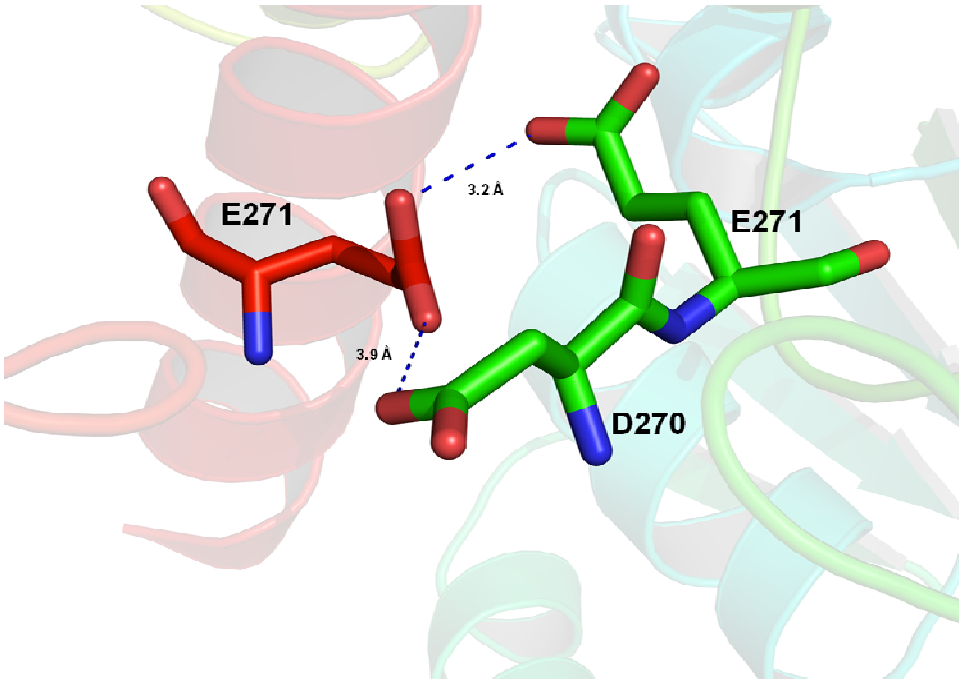


B)

Figure 3.20: A) Stereo ribbon representation of the dimer of DR TrxR. Bound FAD molecules are shown in stick representation. Subunits are colored by domain organization, with the FAD-binding domains in red and cyan and the NAD(P)H-binding domains in green and magenta. B) The HP TrxR crystallographic symmetry dimer (colored red and blue) generated by applying a transformation matrix (found in remark 350 from PDB code: 3ISH) to the monomers across the two-fold of the $P3_221$ space group. The other dimers (non-crystallographic symmetry) are shown cyan and magenta. Crystallographic and non-crystallographic symmetry dimers are equivalent.



A



B

Figure 3.21: HP TrxR and DR TrxR dimer interface. A) Interactions of residue F267 in the dimer interface of HP TrxR. The two subunits are indicated in different colors (red and magenta). B) Interactions of the Ramachandran outlier, residue E271 in the dimer interface of DR TrxR. The two subunits are also indicated in different colours (red and magenta).

3.7.6 Comparison with other Bacterial TrxR Structures

Crystal structures of several oxidized bacterial TrxRs have been determined to date; those discussed in this thesis and those from *M. tuberculosis* and *A. thaliana* (Gustafsson et al, 2007). They all share moderate sequence identity (DR TrxR and HP TrxR show 44% and 37% sequence identity with EC TrxR respectively) and superposition of their three-dimensional structures shows that their overall folds are similar (Figure 3.22). Sequence alignment of the LMW TrxRs with known three-dimensional structure together with the secondary structural elements of DR TrxR (Figure 3.18), shows that residues from the cofactor binding sites, Trx binding sites and dimer interfaces are mostly conserved. Loop regions are relatively less conserved, with some enzymes containing insertions in the loop regions and at the N- and C- termini. Structures determined under reducing conditions have the same fold and same overall conformation (FO). However, the FAD conformation is different between reduced and oxidized structures. The isoalloxazine ring of FAD is flat in the oxidized structure but displays a bent "butterfly" like structure in the reduced structure (Gustafsson et al, 2007). The observed bending angle of reduced FAD varies from structure to structure. The bending angle of reduced FAD for HP TrxR (22º) lies within the range (15–28º) suggested by molecular-orbital calculations for reduced free flavins (Dixon et al, 1979), whereas that of the EC TrxR (34º) lies outside this range (Gustafsson et al, 2007; Lennon et al, 1999).

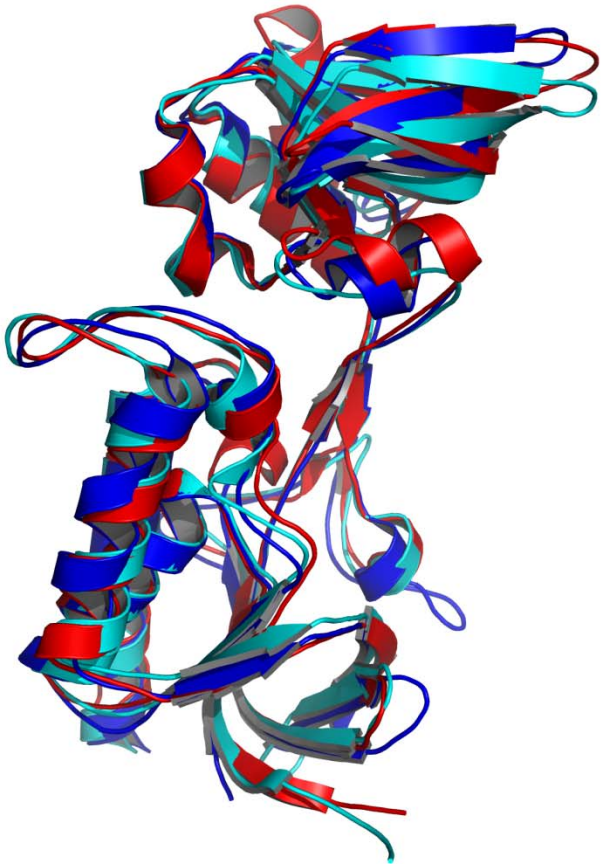


Figure 3.22: Superimposed TrxR monomers (FO conformation) from *E. coli* (red), *D. radiodurans* (cyan), and *H. pylori* (Blue); DR TrxR and HP TrxR structures superpose with the EC TrxR structure with rmsd values of 1.6Å and 1.7Å respectively.

3.8 CONFORMATIONAL STATES OF THIOREDOXIN REDUCTASE

3.8.1 Fluorescence Spectroscopy

As discussed in Section 1.2.4, LMW TrxR undergoes a large conformational change during catalysis. In the FO conformation the redox active disulfide is adjacent to the flavin, whereas in the FR conformation the NADPH binding site is adjacent to the flavin. The two conformations differ by a 66º domain rotation relative to each other. This domain rotation model has previously been successfully investigated by flavin fluorescence spectroscopy of a C138S active-site mutant in EC TrxR. The flavin fluorescence of this mutant is only 7% that of wild-type EC TrxR, presumably due to quenching by S138 in the FO conformation (S138 is close to the flavin in the FO conformation) (Mulrooney & Williams, 1997). An equivalent DR TrxR mutant (C145 Ser active-site mutant), was used to test whether DR TrxR exhibited similar quenching upon mutation of one of the active site cysteines to serine. The results show that DR TrxR behaves in a similar manner to EC TrxR (Figure 3.23) i.e. flavin fluorescence of C145 DR TrxR is much lower compared to the flavin fluorescence of wild-type DR TrxR (~15% of wild-type DR TrxR). These results suggest that domain rotation upon mutation of active site cysteines is probably common to DR TrxR and EC TrxR, although the trigger for domain rotation is not yet known.

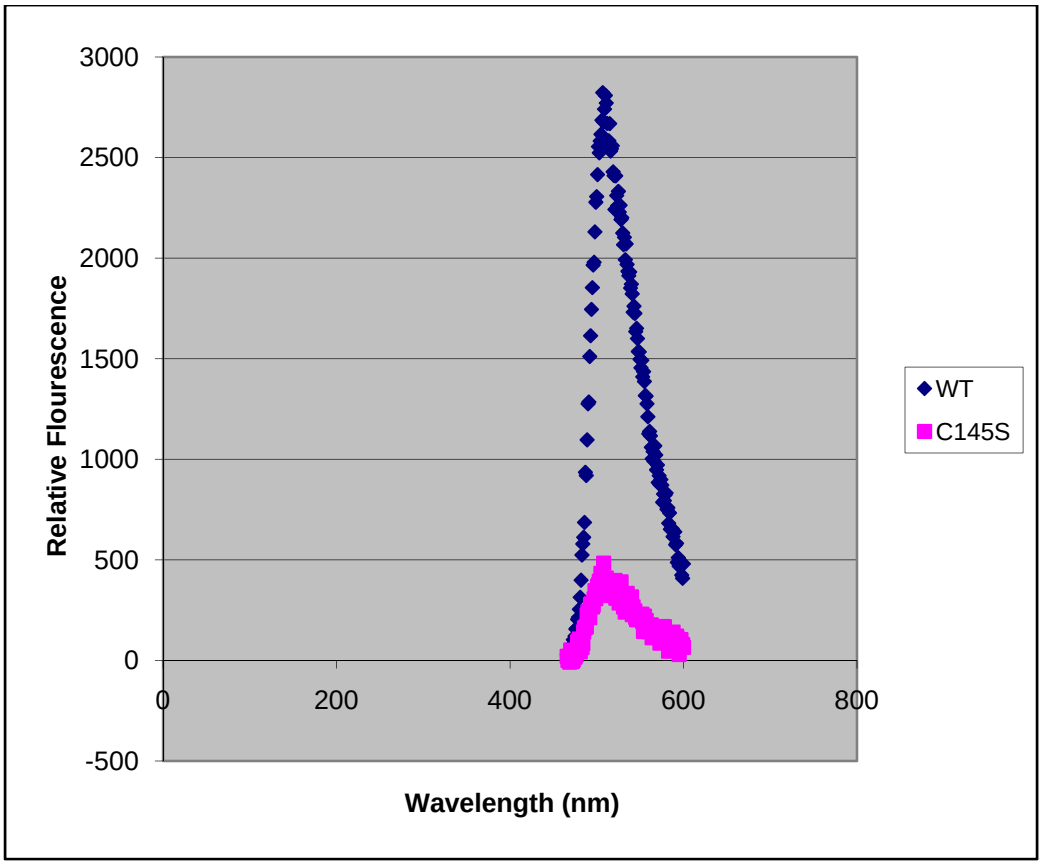
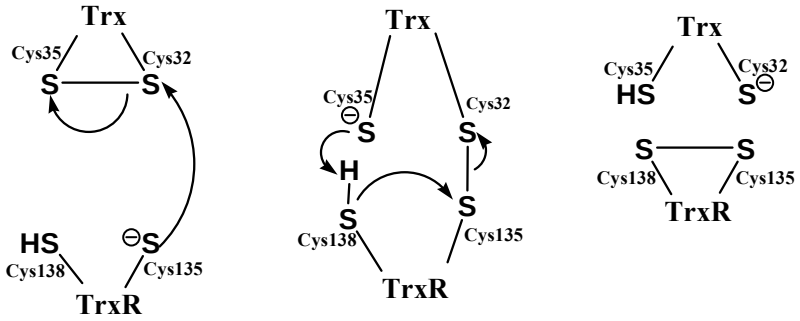


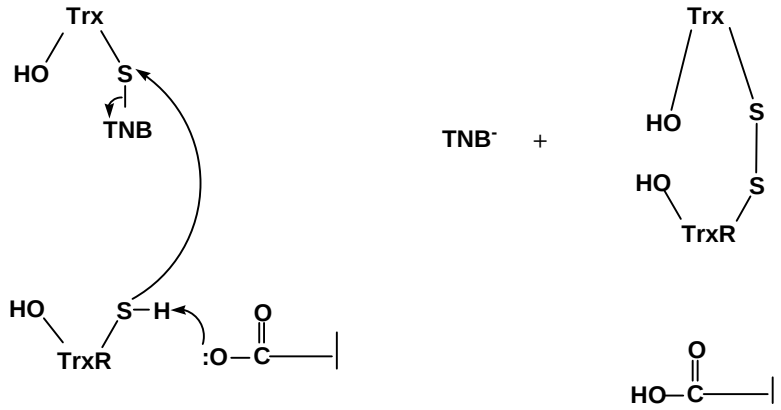
Figure 3.23: Emission spectra of wild type DR TrxR vs C145S mutant. The large decrease in fluorescence suggests that the the FO mutant is predominantly in the FO conformation.

3.8.2 Complex Formation

A covalent complex between EC TrxR and EC Trx has previously been prepared and characterized (Lennon et al, 2000). The crystal structure of the EC TrxR-Trx complex has been determined and shows that EC TrxR in this structure is "trapped" in the FR conformation (Lennon et al, 2000). All other crystal structures of bacterial TrxRs that have been determined to date are in the FO conformation. In order to compare characteristics of homologous and heterologous complexes, a previously described scheme (Wang et al, 1996) was used to generate various homologous and heterologous TrxR-Trx complexes, with the aim of determining their crystal structures. The following mutants were generated by site-directed mutagenesis for the purposes of complex formation: C142S and C145S DR TrxR; C64S and C67S DR Trx1; C28S and C31S HP Trx2. The complexes were formed as described in section 2.5. Trx mutants were first reacted with DTNB to form a mixed Trx-5-thio-2-nitrobenzoic acid (TNB) complex and then the final complexes formed by reacting Trx-TNB complex with TrxR active site mutant (Wang et al., 1996) (Figure 3.24).



A



B

Figure 3.24: Schematic representation of TrxR –Trx complex conformation. The mixed *E. coli* wild type TrxR/Trx disulfide is short-lived (A). Reaction of *E. coli* C135S TrxR with C32S Trx-TNB disulfide results in a stable mixed TrxR/Trx disulfide (B).

An alternative method was previously developed where 4,4'-dithio-dipyridine (DPDS) is reacted with TrxR mutant first to form a mixed TrxR-PDS complex before reaction with the Trx mutant for the formation of the final complex (Veine et al, 1998). Both strategies were used in an attempt to form the complex between DR TrxR and DR Trx1 using various combinations of mutants, because previously it had been shown that complexes can form from whatever

combinations of TrxR and Trx mutants were used. However, the attempts yielded low amounts of DR TrxR- DR Trx1 complex as analyzed by SDS-PAGE. The problem was probably due to the low solubility or low reactivity of DR Trx1. The heterologous system, DR TrxR- HP Trx2 complex, yielded a higher amount of complex as analyzed on an SDS-PAGE (Figure 3.25).

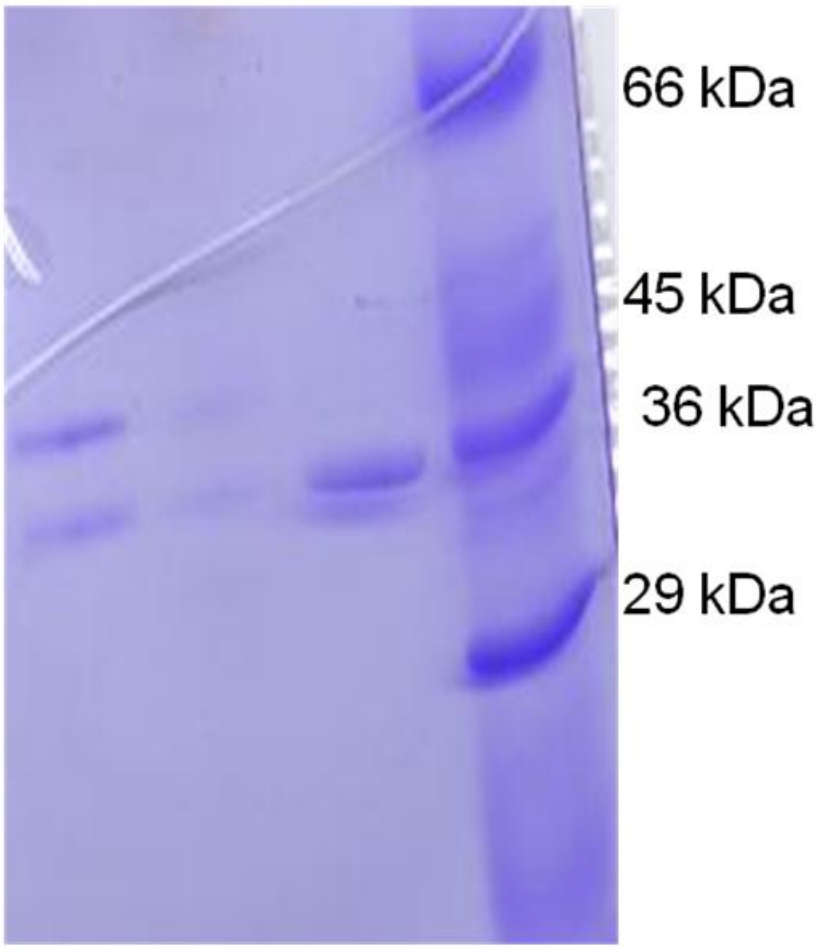


Figure 3.25: SDS-PAGE of the DR TrxR-HP Trx2 complex. Lane 1 shows the Non-reducing SDS-PAGE and lane 2 shows the reducing SDS-PAGE of the DR TrxR-HP Trx2 complex. The expected molecular weights are DR TrxR-HP Trx2 complex (48 kDa), DR TrxR (36 kDa), and HP Trx2 (12 kDa) (too small to be seen on this gel).

3.9 THIOREDOXIN REDUCTASE SPECIES SPECIFICITY

3.9.1 Cross Reactivity of the DR Trx System with the EC Trx System

DR TrxR and DR Trx1 were tested for their ability to cross-react with EC Trx1 (34% sequence identity with DR Trx1) and EC TrxR (38% sequence identity with DR TrxR) respectively, as described in the Materials and Methods section. These cross-reactions are considered heterologous because they involve interactions between proteins of different species. The results are summarized in Table 3.1. Previous studies have shown that TrxRs have a higher affinity for cognate Trxs than for Trxs from other species (Table 3.3) (Arner et al, 1999; Oblong et al, 1993; Oliveira et al, 2010; Pervin, 2006). Consistent with results from other bacterial Trx systems, DR TrxR prefers its cognate Trx (K_m, 2.7 μM) as a substrate over Trxs from different species (EC Trx1 K_m, 10.7 μM) (Table 3.3).

Protein	**K_m for homologous Trx**	**K_m for heterologous Trx**
E. coli TrxR	2.0 μM	20.0 μM (Human Trx)
E. coli TrxR	2.0 μM	35.0 μM (Rat Trx)
A. thaliana TrxR	3.0 μM	80.0 μM (*E. coli* Trx)
D. radiodurans TrxR	2.7 μM	10.7 μM (*E. coli* Trx1)
S. cerevisiae TrxR	1.3 μM	ND (*E. coli* Trx1)
S. cerevisiae TrxR	1.3 μM	ND (Human Trx)

Table 3.3: K_m values of homologous and heterologous Trx systems. Heterologous sytems have higher K_m values compared to homologous systems. The K_m values for interaction between *S. cerevisiae* TrxR and human Trx or EC Trx1 was too high to be determined.

The differences in specificity have been attributed to differences in the size and charge of loop regions of the FAD-binding domains (Oliveira et al, 2010; Zhang et al, 2009) and differences in the charge and shape of Trx-binding surfaces (Gustafsson et al, 2007; Obiero et al, 2010; Oliveira et al, 2010). However, the molecular basis of this phenomenon is still not completely understood (Oliveira et al, 2010). Structures of LMW TrxRs have provided useful clues about the basis of the species specificity phenomenon (Dai et al, 1996; Oliveira et al, 2010). Although the overall structure of LMW TrxRs is highly conserved (FAD and NADPH domains), there are significant differences in the size and type of residues found on the loop regions between prokaryotic and eukaryotic LMW TrxRs. Simulated TrxR-Trx complex models suggest a role for the loop regions in Trx binding specificity in eukaryotic LMW TrxRs.

In contrast, there are no significant differences in loop regions among prokaryotic TrxRs. Moreover, the EC TrxR-Trx complex structure (PDB code: 1F6M) shows that the residues involved in interactions between TrxR and Trx are mainly from the NADPH binding domain and not at the loop regions. The basis of the species specificity within prokaryotic LMW TrxRs has therefore been thought to be due to the subtle differences of the NADPH domain (Gustafsson et al, 2007; Obiero et al, 2010). However, there has not been any systematic study of the Trx-binding residues of prokaryotic LMW TrxRs. This is the first description of the role played by specific residues on the prokaryotic LMW TrxR-Trx interface.

3.9.2 Structural Analysis of the Trx Binding Site

Our kinetic results show that DR TrxR has a lower affinity (4-8 fold higher K_m) for EC Trx1 compared to its own Trx. The structural basis of binding of EC Trx1 to EC TrxR has been elucidated (Lennon et al, 2000). Eight residues (G129, R130, G131, V132, S133, F141, F142 and Y143) on the surface of the NADPH domain of EC TrxR form a groove to which a complementary Trx loop binds. Two of these residues (F141 and F142) fit into a hydrophobic pocket on EC Trx1, making additional interactions with EC Trx1 (Lennon et al, 2000). The two EC TrxR residues that fit into a hydrophobic pocket on EC Trx have corresponding residues in DR TrxR (F148 and 149 respectively) and HP TrxR (F139 and F140). However crystal structures of EC TrxR, DR TrxR, and HP TrxR reveal that these residues poorly superpose and point in different directions (Figure 3.26). The two residues together with R130 constitute the largest proportion of the Trx binding site surface area. It can thus be inferred that variation in the shape of these residues contributes most to species specificity of TrxR-Trx recognition. All the eight residues that form the Trx binding groove on the EC TrxR surface are identical to those of DR TrxR and HP TrxR except the conserved substitution of R130 (EC TrxR) with K137 (DR TrxR) or K128 (HP TrxR). This substitution also contributes to differences in the shape of Trx binding site (Figure 3.26).

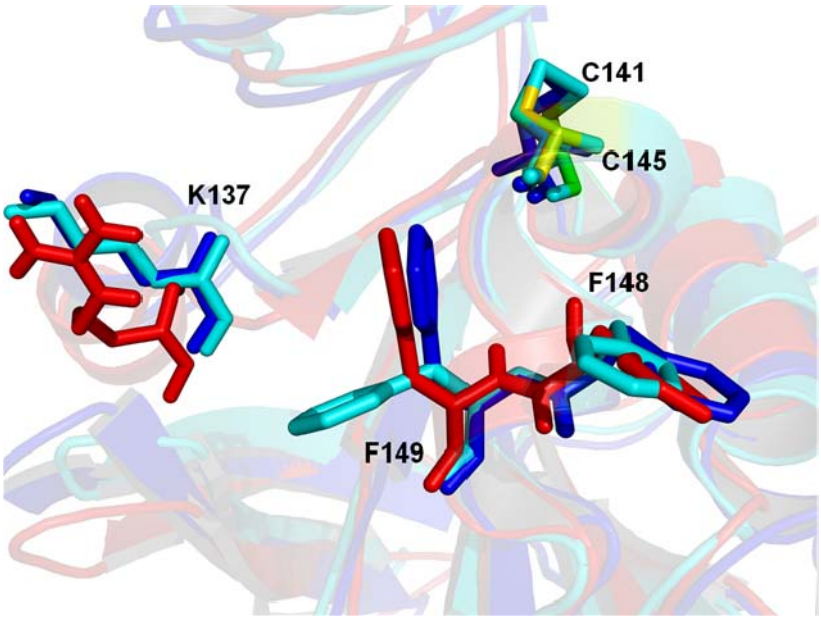


Figure 3.26: Superimposed TrxR structures (FO conformation) from *E. coli* (red), *D. radiodurans* (cyan), and *H. pylori* (Blue) showing some key residues that are important for binding to thioredoxin. Active site disulfide is also shown.

3.9.3 Modeling of TrxR-Trx Complex Structures

A more thorough analysis of shape and complementarity differences between homologous and heterologous TrxR-Trx complexes could be performed if homologous and heterologous complexes were available. Several fundamental properties can be obtained from a crystal structure or homology model of a protein-protein complex including overall size, shape complementarity and electrostatic complementarity of the interface (Jones & Thornton, 1996). These properties can be used to gain understanding of transient protein-protein interactions at a molecular level. The most popular approach of evaluating electrostatic complementarity between protein-protein interfaces is to calculate electrostatic potentials of the interface using the Poisson-Boltzmann equation (Elcock et al, 2001). Shape complementarity can be evaluated by planarity or gap volume index calculations (Jones & Thornton, 1996). A planarity score can be obtained by calculating the rmsd of all interface atoms (within 4Å of each other) from an imaginary plane at the centre of the interface. A lower planarity score suggests greater complementarity. The gap

volume index relates volume of interface cavities to the interface area. It is calculated by the following equation:

GV (Å) = Gap volume between molecules ($Å^3$)/Interface area ($Å^2$) (per complex).

Similar to planarity, a lower gap volume index also suggests better shape complementarity (Bahadur & Zacharias, 2008; Jones & Thornton, 1996).

In the absence of complex structures, previous studies have shown that a reasonable protein-protein complex can be modeled by protein-protein docking from two unbound protein partners (Chen & Weng, 2002). Since the DR TrxR-DR Trx complex structure is currently unavailable, docking was used to generate DR TrxR-DR Trx complex for computational analysis. Using homology models of Trx and TrxR (reduced conformation) and a protein-protein docking software (Z-DOCK), a model of the DR TrxR-DR Trx complex was generated (along with models of other complexes in Table 3.4).

Trx system	Interface area ($Å^2$)	GV Index (Å)
DR TrxR/DR Trx1	**1162**	**0.2 (2.2)**
DR TrxR/EC Trx1	1030	2.8
EC TrxR/EC Trx1	**1160**	**2.5**
EC TrxR/DR Trx1	1259	3.3 (2.9)

Table 3.4: Surface area and gap volume indices. The values in parentheses represent indices from the *D. radiodurans* Trx1 model generated from *R. capsulatus* Trx2 crystal structure (PDB code: 2PPT).

In this study, both electrostatic and shape complementarity were used to study the differences between homologous and heterologous TrxR-Trx interactions. Gap volume index calculations (Table 3.4) suggest that homologous TrxR/Trx systems show more complementarity than

heterologous systems. The gap volume index for EC TrxR/DR Trx complex is higher (3.3Å) than that of EC TrxR/ EC Trx complex (2.5Å); likewise the gap volume index for DR TrxR/ EC Trx complex is higher (2.8Å) than that of DR TrxR/ DR Trx complex (0.2Å). A comparison of the overall surface-charge distribution shows some charge complementarity between Trx and TrxR. The surface surrounding the Trx-TrxR complex interface is both positively and negatively charged in Trx as well as in TrxR. The interface itself is mostly non-polar in both Trx and TrxR. This complementarity is seen in all the Trx systems analyzed (*E. coli*, *D. radiodurans,* and *H. pylori*). The surface around the Trx binding site of EC TrxR was the most negative and that of DR TrxR was the most positive (Figure 3.27). This difference in surface charge may also play a role (although probably small role compared to shape) in the species specific recognition.

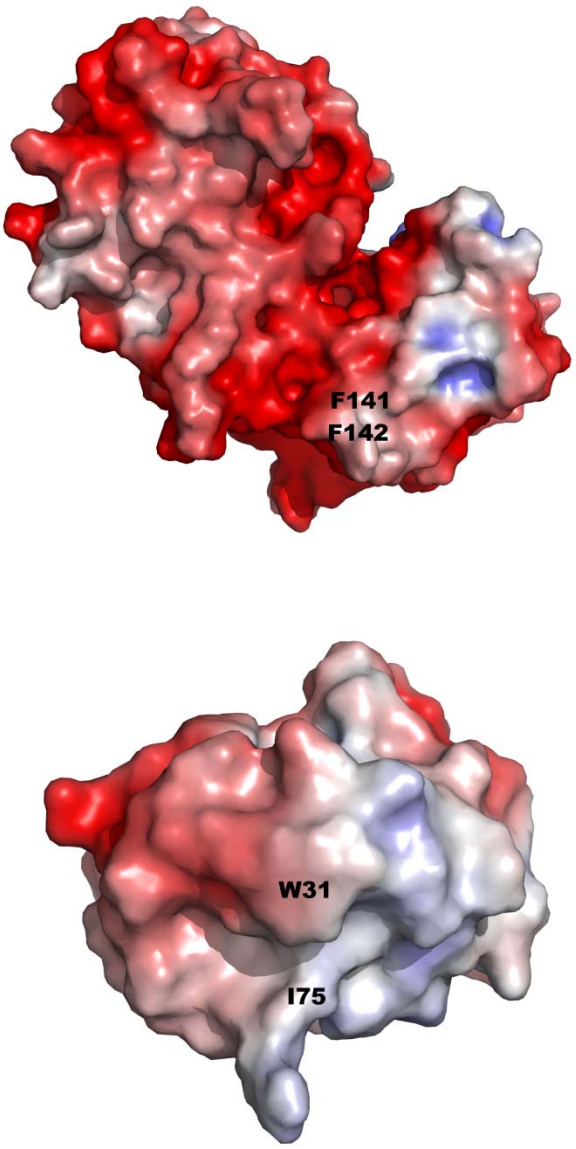


Figure 3.27 A: Electrostatic surface potentials of *E. coli* TrxR (top) and *E. coli* Trx1 (bottom)

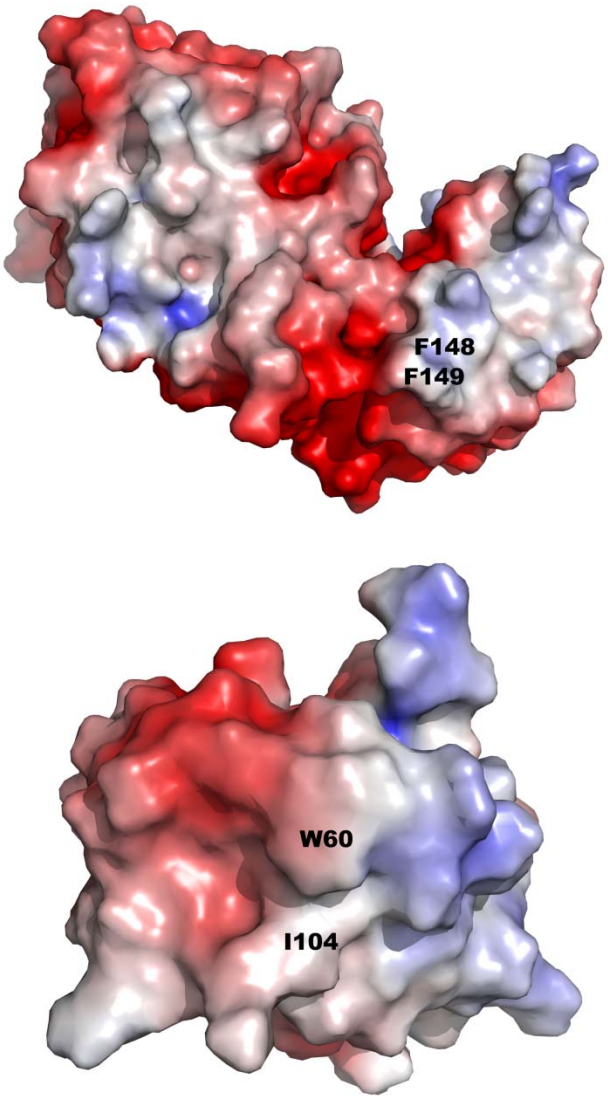


Figure 3.27 B: Electrostatic surface potentials of *D. radiodurans* TrxR (top) and *D. radiodurans* Trx1 (bottom)

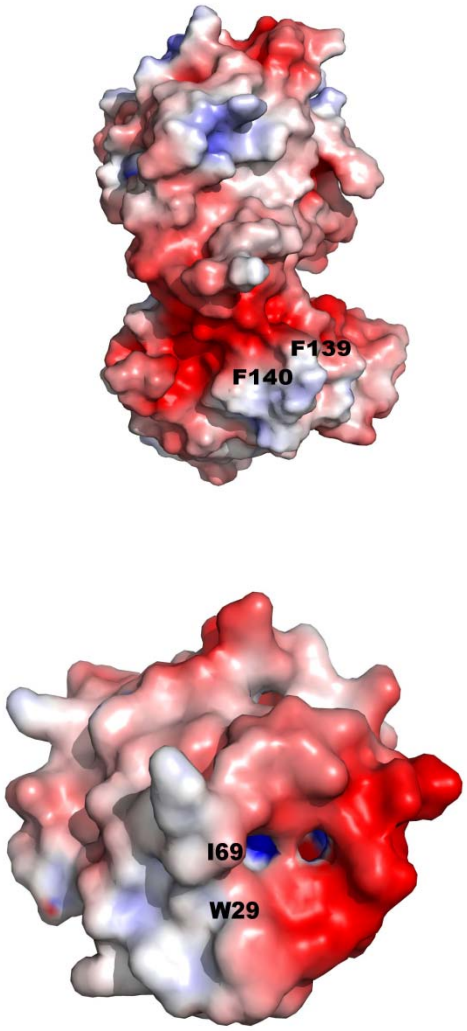


Figure 3.27 C: Electrostatic surface potentials of HP TrxR (top) and HP Trx1 (bottom). Electrostatic potentials for all the figures (A,B, and C) were calculated by APBS software (Baker et al, 2001b). Positively and negatively charged areas are colored red and blue and contoured at -5 k_BTe^{-1} and +5 k_BTe^{-1} respectively. Key residues of the TrxR-Trx complex interface are shown on the diagram. The overall shape of Trx-binding site differs between the three TrxRs, but there are only small differences in the electrostatic potentials. The figures were generated with PYMOL using EC TrxR, EC Trx, DR TrxR, and HP TrxR coordinates (PDB codes 1TDE, 2TRX, 2Q7V, and 3ISH respectively). The DR Trx1 and HP Trx1 figure were generated using the DR Trx and HP Trx1 homology models.

3.9.4 Identification of Hot Spot Residues

Previous research on protein-protein interactions has shown that only a small fraction of the interface residues (hot spot residues) are responsible for the majority of interface stability (Moreira et al, 2007). Alanine scanning mutagenesis is the most widely used technique to identify hot spot residues. It was first used to map hot spots of binding energy in the human growth hormone-receptor interface (Clackson & Wells, 1995). Although slow and labor intensive, it has since been used to map binding hot spots for several protein-protein interfaces (Bogan & Thorn, 1998; DeLano, 2002; Moreira et al, 2007; Thorn & Bogan, 2001). Alanine scanning measures the effect of the deletion of an amino acid side chain beyond the β-carbon atom on the affinity of a protein-protein complex. Glycine substitutions would be ideal for measuring the effect of deletion of all the amino acid side chain atoms but glycine can introduce additional conformational freedom into the protein backbone, therefore it is rarely used (Moreira et al, 2007). Because experimental alanine scanning mutagenesis is slow and labor intensive, a computational protocol (computational alanine scanning mutagenesis) has been developed to map protein-protein interfaces (Kortemme et al, 2004b). The protocol uses a simple free energy function to calculate the effects of alanine mutations on the binding free energy of a protein-protein complex. This technique correctly predicted 80% of protein-protein interface hot spot residues, when compared with experimentally observed change in binding free energy upon alanine substitution.

Using this technique with the Robetta server (http://www.robetta.org/submit.jsp), three EC TrxR residues of TrxR-Trx complex (PDB code: 16FM) interface were predicted to be hot spot residues (Table 3.5). Previously it has been suggested that binding hot spots are those residues which show a $\Delta\Delta G$ larger than 1 kcal/mol (Kortemme & Baker, 2002). Only three (F81, F141

and F142) of the 22 residues had a predicted value of greater than 1 kcal/mol. Although its predicted ΔΔG was less than 1 kcal/mol, visual inspection of the EC TrxR-EC Trx interface suggested a significant role for R130. Five residues form EC TrxR (D99, N100, G129, R130 and A237) and four residues from EC Trx (W31, E44, S95 and R73) participate in hydrogen bonding across the TrxR-Trx interface. Of the EC TrxR residues that form hydrogen bonds across the TrxR-Trx interface, Arg 130 contributes most to interface stability. The side chain of this residue participates in four hydrogen bonds with amino acids from Trx. It was thus deemed to also play a significant role in interface stability. Equivalent residues (to the four EC TrxR residues) from DR TrxR were then individually mutated to alanine and tested for affinity towards Trx.

Residue #	Amino acid type EC TrxR	Amino acid type DR TrxR	ΔΔG (complex) kcal/mol
37	M	G	0.69
38	E	M	0.00
39	K	P	0.04
44	T	A	0.00
45	T	W	0.08
81	F	M	2.01
82	D	D	-0.11
83	H	E	0.24
99	D	Y	0.30
100	N	N	0.40
130	R	K	0.78
132	V	V	-0.01
133	S	S	0.41
135	C	C	-0.02
137	T	T	0.08
138	C	C	-0.59
139	D	D	0.11
141	F	F	1.70
142	F	F	3.27
143	Y	Y	0.56
215	M	-	0.13
217	V	V	-0.05

Table 3.5: Computational alanine scanning mutagenesis results from Robetta server. Hot spot residues are highlighted in red. Hot-spot residues can be defined operationally as those for which alanine mutations have destabilizing effects on ΔΔG of more than 1 kcal/mol. Interface residues are defined as those residues with atoms within 4Å of the other partner.

3.9.5 Kinetic Analysis of Hot Spot Residues

Equivalent residues (to the four EC TrxR residues highlighted in Table 3.5) from DR TrxR were individually mutated to alanine and tested for affinity towards Trx. All the DR TrxR alanine mutants showed a significant reduction in affinity compared to wild type DR TrxR when reacted with DR Trx1. The F148A TrxR mutant showed the greatest increase in K_m (20 fold), followed by the F149A TrxR mutant (10 fold) and M84A TrxR mutant (8 fold). The K137A TrxR mutant showed the smallest increase in K_m (4 fold), when compared with homologous DR Trx system. A similar trend was observed when the same DR TrxR mutants were cross-reacted with EC Trx1. The F149A TrxR mutant showed the greatest increase in K_m (10 fold), followed by F148A (6 fold) and then the K137A TrxR mutant K_m (2 fold), when compared with K_m values of the wild type heterologous cross-reaction assay (DR TrxR/EC Trx1). However, there was no significant change in the K_m values M84A mutant and the K137R mutants in the cross-reaction assays. Instead both DR TrxR mutants showed an increase in the specificity constant. The results are summarized in Table 3.6.

Assay	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
DR TrxR (wt)/DR Trx1	**2.8 ± 0.4**	**9.4 ± 1.0**	**3.4×10^6**
DR TrxR M84A/DR Trx1	47.8 ± 2.7	3.2 ± 0.2	5.4×10^4
DR TrxR K137A/DR Trx1	10.4 ± 0.9	8.9 ± 0.3	8.6×10^5
DR TrxR F148A/DR Trx1	61.4 ± 6.7	11.0 ± 0.5	1.8×10^5
DR TrxR F149A/DR Trx1	26.9 ± 3.3	2.2 ± 0.1	8.1×10^4
DR TrxR K137R/DR Trx1	5.4 ± 0.4	7.2 ± 0.2	1.3×10^6

A)

Assay	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
DR TrxR (wt)/EC Trx1	**10.4 ± 1.0**	**2.2 ± 0.4**	**2.1×10^5**
DR TrxR M84A/EC Trx1	6.5 ± 0.1	11.2 ± 1.0	1.7×10^6
DR TrxR K137A/EC Trx1	17.7 ± 4.1	7.3 ± 0.5	4.1×10^5
DR TrxR F148A/EC Trx1	60.3 ± 4.1	6.8 ± 0.3	1.1×10^5
DR TrxR F149A/EC Trx1	97.2 ± 7.8	2.0 ± 0.1	2.0×10^4
DR TrxR K137R/EC Trx1	8.3 ± 0.3	10.2 ± 0.3	1.2×10^6

B)

Table 3.6: Kinetic parameters Trx binding site mutants of DR TrxR. Most alanine mutants showed an increase in K_m and a decrease in specificity constants (k_{cat}/K_m) when reacted with DR Trx1 (A) and cross-reacted with EC Trx1 (B). The M84A mutant showed an increase in specificity constant when cross-reacted with EC Trx1 compared to DR TrxR (wt)/EC Trx1.

The exact *in vivo* concentration of Trx is not known. As a result there are widely varying estimates of Trx concentrations (1-100μM) (Gromer et al, 2004). Regardless of the exact Trx concentrations *in vivo* the change in affinity due to the substitution of the four residues with alanine was large enough that there would be no meaningful interaction expected *in vivo* without the four residues. Since mutational effects have been shown to be generally additive (Qin et al, 1995; Wells, 1990), changing the four residues to alanine would be expected to lower the affinity of TrxR for Trx by more than 10 000 fold. These results therefore suggest that only four residues are responsible for the majority of the TrxR-Trx interface stability. Of the four residues, two are invariant (F148, F149) across bacterial TrxR and two others have conserved substitutions (M84, K137) (Figure 3.18). Together these four residues can be considered to be the binding epitope of the TrxR-Trx interface.

Mapping specific residues at protein-protein interfaces that contribute to the specificity and strength of protein-protein interactions has been a subject of intense research in the last few years. Alanine scanning mutagenesis has enabled detection of hot spot residues and demonstrated that most of the binding strength at protein-protein interfaces is due to hot spot residues (a small subset of interface residues) (Moreira et al, 2007). These results show that the TrxR-Trx interface is similar to other protein-protein interfaces, where only a small number of interface residues are responsible for most of the binding affinity and specificity. In addition, some binding hot spots interact with a variety of partners (DeLano et al, 2000); therefore, for proteins that interact with several partners like Trx, identifying one binding hot spot may be useful in studying interactions with other protein partners (Qin et al, 1995).

3.9.6 Comparison of Calculated TrxR-Trx Interaction Energies versus Affinities

For cross validation purposes, experimentally observed changes in affinities were correlated to computationally calculated interaction energies. Two DR TrxR-DR Trx complex models were constructed for computational characterization using DR TrxR and two DR Trx models (see section 2.11). The two DR Trx models were constructed using EC Trx1 crystal structure and *R. capsulatus* Trx2 crystal structure. However, only the DR Trx model generated from *R. capsulatus* Trx2 crystal structure was used for calculation of interaction energies because the sequence of DR Trx1 suggests that its closer to Trx2 than Trx1 (DR Trx1 contains extra N-terminal cysteines that are usually found only in Trx2 as seen in Figure 3.4). In addition, DR TrxR-DR Trx complex model generated using the DR Trx model constructed using coordinates of *R. capsulatus* Trx2 crystal structure showed reasonable gap volume index values (Table 3.4).

A comparison was done between the calculated interaction energies and experimentally observed affinities using the aforementioned DR TrxR-DR Trx complex model. The correlation coefficient value was found to be 60% (Figure 3.28) for calculated interaction energy. Consistent with results from previous studies, these results show that parameters like interaction energy can be useful in predicting and ranking hot spot residues of protein interfaces (Khoury et al, 2009; Kortemme et al, 2004b). Although not completely predictive, these results suggest that computational alanine scanning mutagenesis is a useful alternative to the labor-intensive alanine scanning mutagenesis analysis (for identification of hot spot residues and redesign of enzymes with new specificities).

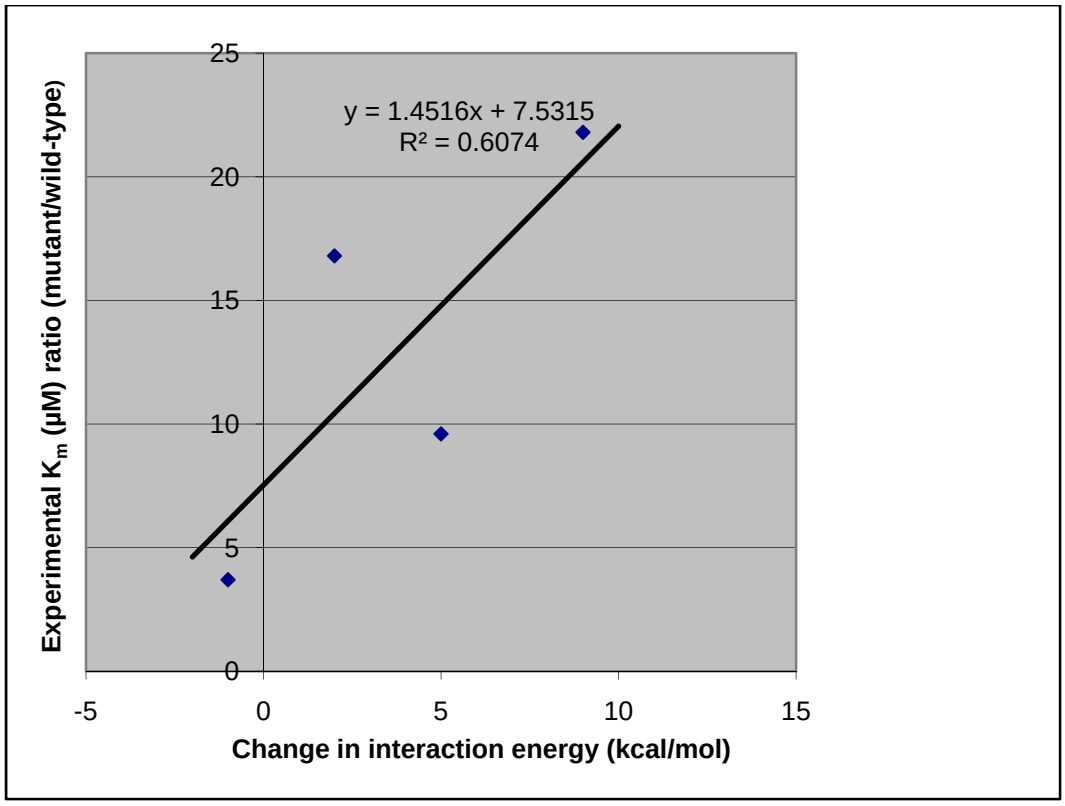


Figure 3.28: Plot of experimental K_m ratios (mutant/wild-type) vs. calculated changes in interaction energies (AMBER) for the DR TrxR mutants described in this study. Interaction energies showed reasonable correlation (R^2 = 60%) with the ratio (mutant/wild-type) of experimental K_m values.

3.9.7 Design of DR TrxR for Altered Trx Specificity

These computational predictions and experimental results suggest that two invariant residues (F148, F149) are responsible for the majority of TrxR affinity for Trx and two others with conserved substitutions are responsible for the majority of specificity (M84, K137). Together the four residues can be considered as the binding epitope of the TrxR-Trx interface. Replacement of DR TrxR M84 with alanine resulted in a large increase in K_m for the homologous assay (DR TrxR M84A/DR Trx1) but decrease in K_m for the heterologous assay (DR TrxR M84A/EC Trx1) (Table 3.6). These results implicated M84 as one of the key residues in determination of species specificity. The DR TrxR K137R mutant also showed an increase in the K_m value for the homologous assay (DR TrxR K137R/DR Trx1) and a slight decrease in the K_m value for the heterologous assay (DR TrxR K137R /EC Trx1) (Table 3.6). Therefore K137 was also implicated in species specificity.

Our working hypothesis was that if all the residues of the functional epitope of DR TrxR were changed to match those of EC TrxR, then DR TrxR would be changed to have a higher affinity for EC Trx than for its own Trx. We therefore made a double DR TrxR mutant (M84F, K137R) and tested its affinity for DR Trx and EC Trx. The results (Table 3.7) are consistent with our predictions; the DR TrxR double mutant had a much higher affinity for EC Trx compared to its own Trx. It is worth emphasizing that, although the four residues are crucial in TrxR-Trx interface stability, they are insufficient in altering the specificity to match or surpass that of the wild-type Trx systems. This suggests a role for other factors such as overall shape and charge of the Trx-binding site in determining species specificity as previously suggested (Gustafsson et al, 2007; Obiero et al, 2010).

Assay	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
DR TrxR (wt)/DR Trx1	**2.8 ± 0.4**	**9.4 ± 1.0**	**3.4×10^6**
DR TrxR (wt)/EC Trx1	**10.4 ± 1.0**	**2.2 ± 0.4**	**2.1×10^5**
DR TrxR M84F,K173R/EC Trx1	4.7 ± 0.8	1.3 ± 0.2	2.8×10^5
DR TrxR M84F,K173R/DR Trx1	16.5 ± 1.5	0.9 ± 0.1	5.4×10^4

Table 3.7: Kinetic parameters DR TrxR double mutant (M84F, K137R). The double mutant has a higher K_m value for DR Trx1 double mutant than EC Trx1, demonstrating a switch in specificity.

CHAPTER FOUR: CONCLUSIONS AND PERSPECTIVES

4.1 CONCLUSIONS

Intermolecular forces are the major determinants in many biological processes. These forces play an important role in many processes that require specific recognition such as signal transduction, regulation, immune recognition and adhesion. They are also important in other biological processes such as nucleic acid folding, protein folding and catalysis. A major research goal over the years has been to develop an approach which can computationally determine free energies for non-covalent interactions in complex molecules (Massova & Kollman, 1999). As a result various approaches have been developed that can calculate the effect of mutations on the binding free energy of specific protein-protein complexes (Guerois et al, 2002; Huo et al, 2002; Kortemme & Baker, 2004; Massova & Kollman, 1999; Sharp, 1998). These computational protocols use different energy functions, but there is good agreement among them (Kortemme et al., 2004). Previous studies have shown that proteins can be computationally designed to bind new ligands (Looger et al, 2003), proteins (Kortemme & Baker, 2004), and nucleic acids (Ashworth et al, 2006). Using the EC TrxR-Trx complex structure as a guide and site-directed mutagenesis, the DR TrxR's substrate specificity has been successfully altered. First, the residues that enable bacterial TrxR to recognize and discriminate between cognate Trxs and Trxs from different species were identified (Figure 4.1). Then the role of these residues was validated by the successful design of TrxR mutants that exhibited altered specificity for Trx. The DR TrxR double mutant (M84F, K137R) showed a higher affinity for EC Trx than DR Trx.

Understanding the basis of TrxR-Trx interactions provides a framework for design of Trx systems with desirable efficiency and selectivity for application in pharmaceutical and biotechnological industries. In particular, understanding the basis of the presumed cross-

interaction between secreted *H. pylori* Trxs and host proteins (mucins and immunoglobulins) has a huge therapeutic potential. HP Trx is thought to mediate catalytic reduction of human immunoglobulins and hence facilitates immune evasion. It has also been proposed that *H. pylori* gains access to the impenetrable gastric mucous layer by using secreted HP Trx to reduce the disulfide bonds present in the cysteine-rich mucin regions that are responsible for cross-linking mucin monomers. Disruption of secreted HP Trx-host protein interaction may result in restoration of the viscoelastic and hydrophobic protective properties of mucus (Windle et al., 2000).

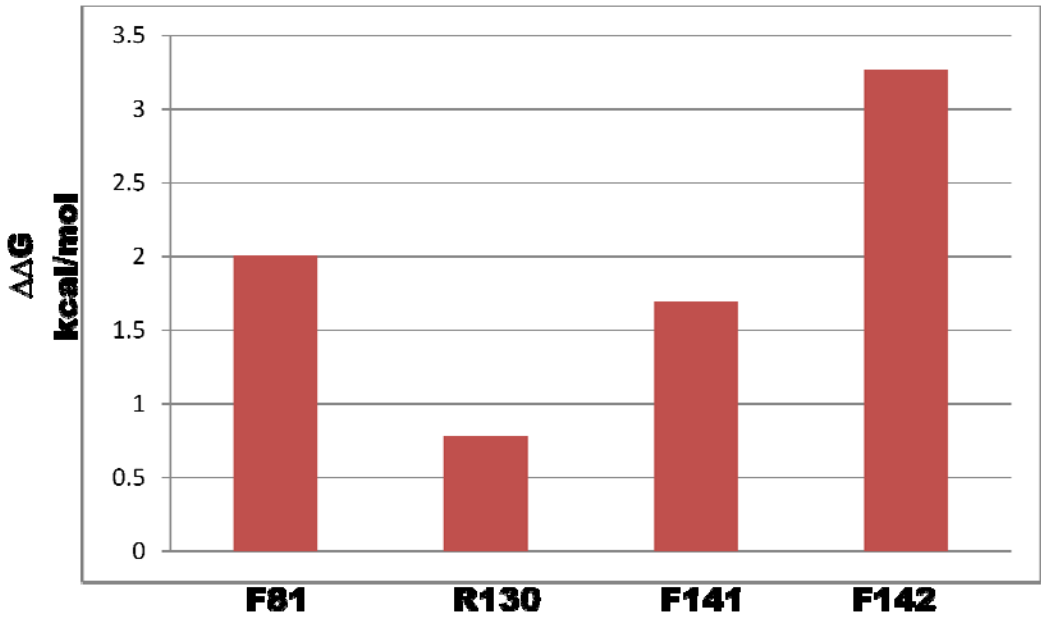


Figure 4.1: Computational alanine scanning mutagenesis results from Roberta Server using (EC TrxR-Trx complex structure) showing important residues for interface stability from EC TrxR**.** Calculated change in ΔΔG upon mutation of Trx-binding residues to alanine. The higher value of ΔΔG signifies a higher contribution of that residue to interface stability. Equivalent residues from DR TrxR were identified and experimentally validated as important for interface stability.

4.2 FUTURE DIRECTIONS

Trx-TrxR interactions are side chain mediated protein-protein interactions. We have demonstrated that a computationally designed DR TrxR mutant can switch specificity for Trx. It would be informative to prepare DR Trx (or any other bacterial Trx) mutants that have altered specificity, using similar computational and experimental strategies. Using the Robetta server the following *E. coli* Trx residues are predicted to contribute most to interface stability: W31, M37, R73, and I75 (Table 4.1). Of these, only one is invariant W31 (DR Trx W60). The rest are not conserved or contain conserved substitutions (Figure 4.2). Therefore an opportunity exists to alter specificity through site-directed mutagenesis. The ultimate goal is to design Trx systems with desirable specificities and efficiencies for application in various industries. Catalytic efficiency of the fastest enzymes (e.g. catalase) is more than 1000 fold higher than bacterial TrxR; therefore an opportunity exists to design TrxR systems with much higher catalytic efficiencies than the wild type enzymes. Along with other properties such as heat stability, enzymes so designed will be more efficient in industrial and pharmaceutical applications.

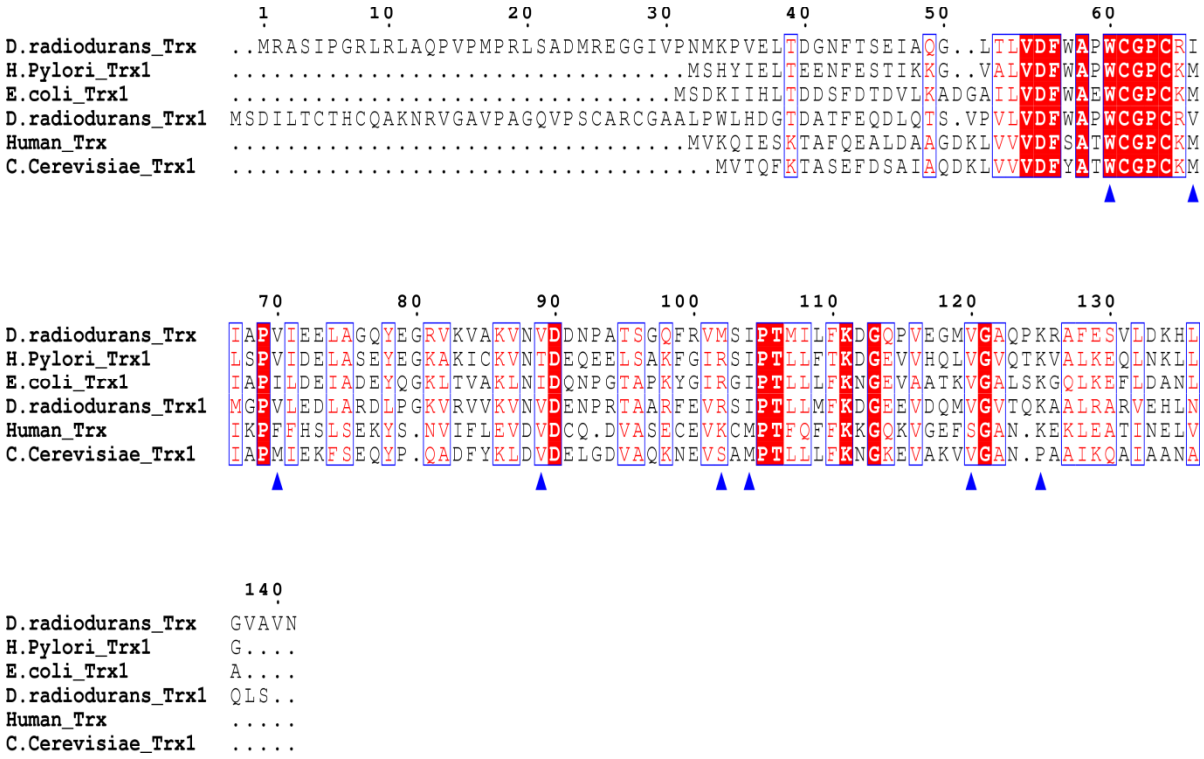


Figure 4.2: Amino-acid sequence alignments of Trxs from various prokaryotic and eukaryotic sources. Invariant residues are highlighted by a red background. TrxR-binding residues are underline by blue triangles.

Residue #	Amino acid type E .coli Trx1	Amino acid type D. radiodurans Trx1	ΔΔG (Complex) kcal/mol
31	W	W	2.74
32	A	A	-0.57
37	M	V	1.04
41	I	V	0.36
44	E	E	0.05
60	I	V	0.67
72	I	V	0.11
73	R	R	2.54
75	I	V	1.10
77	T	T	0.00
91	V	V	0.47
95	S	Q	-0.01
96	K	K	0.34
98	Q	A	0.06

Table 4.1: Computational alanine scanning mutagenesis results from Robetta server using the EC TrxR-Trx complex (PDB code: 1F6M). Hot spot residues are highlighted in blue. Hot-spot residues can be defined operationally as those for which alanine mutations have destabilizing effects on ΔΔG of more than 1 kcal/mol.

APPENDIX

A) CHARACTERIZATION OF THIOREDOXIN-THIOREDOXIN REDUCTASE INTERACTIONS BY ISOTHERMAL TITRATION

A1: Introduction

Calorimetry is the principal source of thermodynamic information, because virtually all chemical and biological processes are accompanied either by liberation (exothermic reaction) or absorption (endothermic reaction) of heat. Isothermal titration calorimetry (ITC) is a calorimetric method designed to measure small amounts of heats generated by interactions between two reactants in a liquid solution at constant temperature (Pierce et al, 1999; Saboury, 2004). An ITC experiment can determine the affinity constant (*Ka*) between two interacting molecules, and the associated thermodynamic parameters: the change in enthalpy (ΔH), the change in entropy (ΔS), the change of Gibbs free energy (ΔG), and the reaction stoichiometric index (*n*). In addition, if the ITC experiments are performed at different temperatures the variation in heat capacity (ΔCp) can be obtained as well (Bundle & Sigurskjold, 1994; Pierce et al, 1999).

ITC is universally applicable to all reactions regardless of the chemical nature or size of the interacting components. Moreover there is no chemical modification or immobilization needed for an ITC experiment. It has thus been widely used to study macromolecule interactions (protein-ligand, protein-protein, and protein-nucleic acid interactions) (Cooper, 1999; Pierce et al, 1999; Weber & Salemme, 2003). Figure A1 below shows the basic design of an ITC instrument.

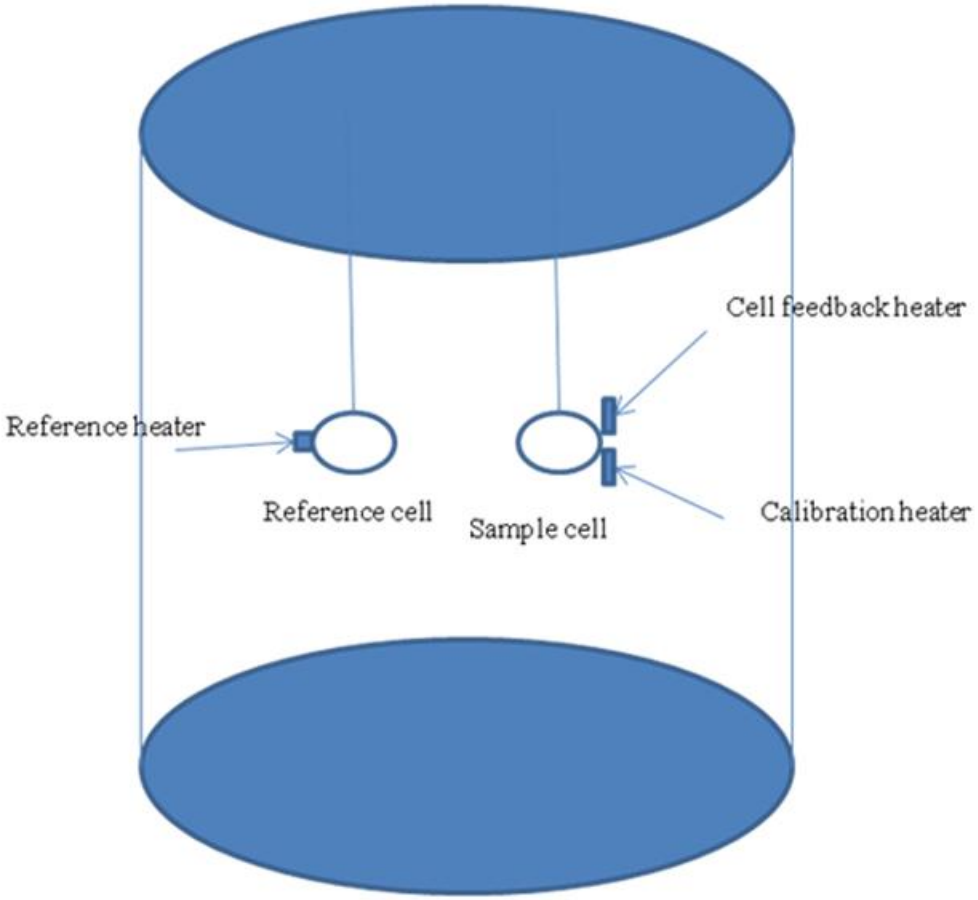


Figure A1: Basic design of an ITC instrument. There are two cells in the adiabatic jacket: the reference cell and the sample cell. A small continuous power is applied by the heater on the reference cell. Injection of ligand into the sample cell leads to interaction between the ligand and macromolecule. Depending on the nature of the interaction, heat is either taken up or evolved. Thermopile/thermocouple detectors sense temperature differences between the sample and reference cells. The feedback circuit then either increases or decreases power to the sample cell to maintain equal temperature with the reference cell. The heat per unit time (μJ/sec) supplied to the sample cell is the observable signal in an ITC experiment and is proportional to the heat evolved on binding of a ligand to a macromolecule (Pierce et al, 1999).

In a typical ITC experiment, the sample cell is filled with a macromolecule in an appropriate buffer and stirred to ensure good mixing. The reference cell is filled with the same buffer used for the sample cell. For measurements of binding constants, concentrations of macromolecule and ligand should be accurately determined. In addition the final concentrations desired should be carefully selected. A good approximation is to make the final macromolecule concentration at least 10 times the expected binding constant, and the final ligand concentration to be two times

the concentration of the macromolecule. Choice of buffer is also critical when planning an ITC experiment. Buffers with low enthalpy of ionization are desired e.g. phosphate buffers (Maley & Hassler, 2004; Pierce et al, 1999). Several small volumes of the ligand are injected and changes in thermal energy due to association are detected and recorded. The raw data obtained is power versus time for each injection. A sample thermogram is shown in Figure A2.

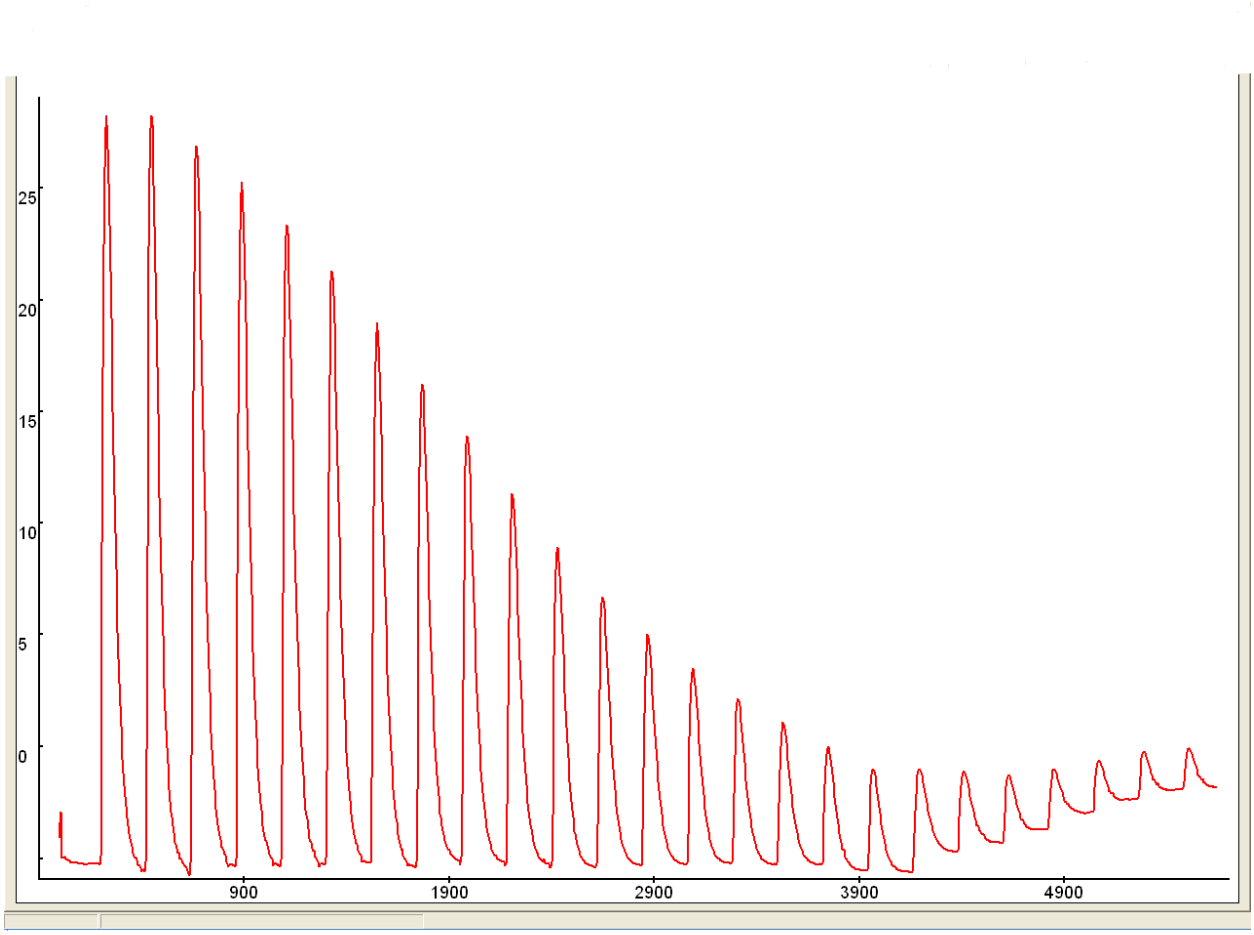


Figure A2: An ITC thermogram of titration of RNase with 3' CMP. Raw data obtained for 25 injections of 4 µL of 3' CMP solution (13mM) into a sample cell containing 1300 µL RNase solution at 0.6mM. The reaction was carried out in a filtered and degassed 200mM potassium phosphate buffer pH, 5.5.

A2: Materials and Methods

The ITC measurements were performed on a CSC ITC-4200 calorimeter at the SSSC. Purified EC TrxR and EC Trx were dialyzed against the reaction buffer (100 mM potassium phosphate buffer, pH 7.5) at 4°C and degassed by stirring under vacuum for 10 minutes just before titration. Various EC Trx concentrations (1000-2000 μM) were injected (15-20 injections of 10μL using a 250μL Hamilton syringe) into sample cell containing 1.30 mL of EC TrxR solution (10 μM). The sample cell was continuously stirred at 300 rpm with 3 minutes between injections.

A3: Results and Discussion

Currently two spectrophometric assays (insulin reduction and DTNB reduction assays) are popularly used two characterize interactions between TrxR and Trx. The two assays are relatively easy and cheap, however, they have some limitations; for example, DTNB inactivates rat TrxR in the presence of Trx (Luthman & Holmgren, 1982), and HP TrxR reduces insulin even in the absence of thioredoxin (Pervin, 2006). Moreover, evaluating protein-protein interactions using Michealis-Menten kinetics involves the assumption that the rate of product formation and release is much slower than that of substrate binding ($k_2 << k_{-1}$) as discussed in section 1.3.1. It has been shown that the rate of product formation (domain rotation during catalysis) is the slowest step in the reaction (Lennon & Williams, 1997), however, it is not clear if the difference is large enough for $K_m \approx k_{-1}/k_1$, the dissociation constant. ITC thus offers an attractive alternative to spectrophotometric assays in characterizing the TrxR-Trx interactions.

Using ITC, it is possible to characterize interactions between TrxR and Trx under reducing (presence of NADPH) and non-reducing conditions. It is also possible to characterize the interaction between the wild-type enzyme and active-site mutants. The EC TrxR C138S active

site mutant is predicted to be predominantly in the FO conformation, whereas the wild type enzyme has been shown to be predominantly in the FR conformation. It would be interesting to see if there is any interaction with the FO favoring mutant, even though the active site disulfide is buried and unavailable for catalysis. Although it requires a large amount of samples, ITC thus offers huge advantages over spectrophotometric assays in characterization of TrxR-Trx interactions.

A unitless parameter, the Wiseman constant, has been widely used to predict the success of an ITC experiment. This parameter can be calculated by the equation below:

$$c = [M]K_a$$

Where [M] is the concentration of the macromolecule and K_a is the association constant. The best thermodynamic data is obtained by those systems with a Wiseman constant of $10 \leq c \leq 500$ (Turnbull & Daranas, 2003; Wiseman et al, 1989). For the EC Trx system (assuming the K_d to be approximately the same as the K_m), a macromolecule concentration of 10μM would give a Wiseman constant of 5 (close to the recommended range). The ITC experiment was carried out as described in section A2. At this concentration of TrxR there was a low signal to noise ratio. Moreover, it was not possible to reach saturation because of an endothermic to exothermic conversion at approximately a 1:1 TrxR:Trx stoichiometry, possibly because of an exothermic conformational change (Figure A3).

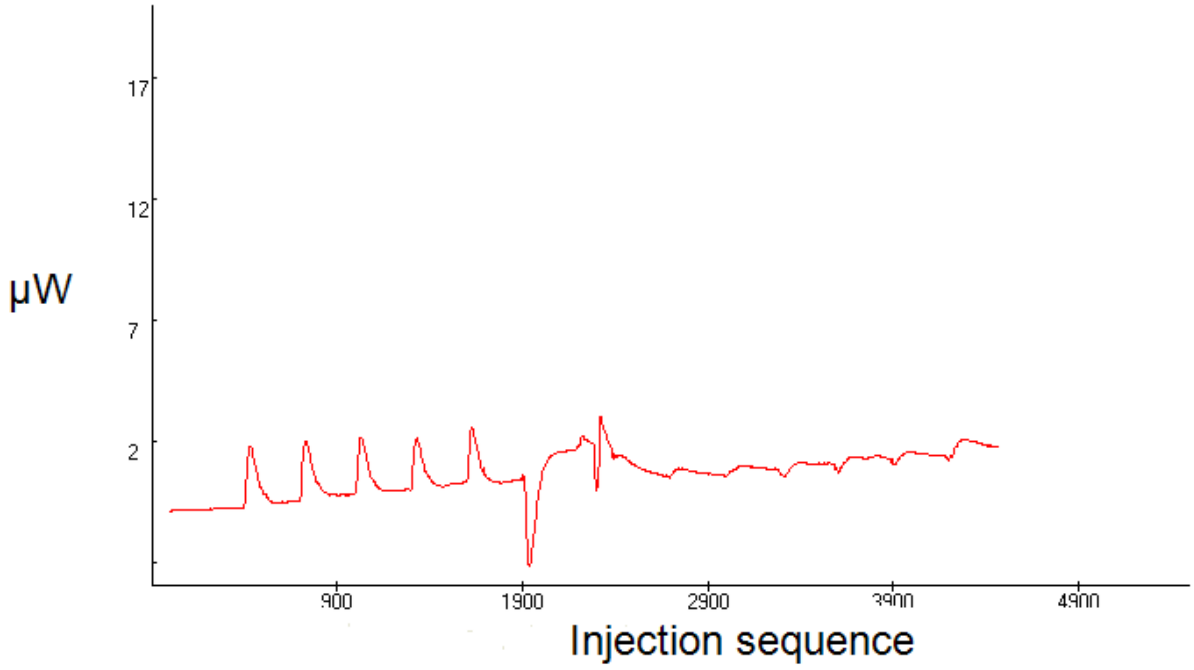


Figure A3: An ITC thermogram of titration of EC TrxR with EC Trx1. Raw data obtained for 15 injections of 10 µL of EC Trx solution (2.0 mM) into a sample cell containing 1300 µL EC TrxR solution at 10 µM. The reaction was carried out in a filtered and degassed 100mM potassium phosphate buffer, pH 7.5.

B) CRYSTALLIZATION OF HP TRX1 AND HP TRX2

B1: Introduction

As discussed in section 1.1.4, *H. pylori* contains two Trxs. The two Trxs show relatively low primary sequence identity to one another (35% sequence identity). Trx1 has the highly conserved Trx active site motif, CGPC, found in most Trxs whereas Trx2 has an unusual motif, CPDC, not found in other Trxs. In addition, Trx2 has an unusually high pI (8.8) compared to Trx1 (pI 5.3). These differences suggest that the two Trxs are functionally dissimilar; which is perhaps an essential divergence in roles because of the absence of glutaredoxin system in *H. pylori* (Baker et al, 2001a). Kinetic characterization (Table B1) also suggests some differences. HP Trx1 has a much higher catalytic efficiency compared to HP Trx2 (Table B1) (Pervin, 2006). Determining the structure of the two Trxs was attempted in order to thoroughly understand how they differ and how they function.

Substrate	**K_m (µM)**	**V_{max} (µM/min)**	**k_{cat} (s^{-1})**	**k_{cat} / K_m ($M^{-1}s^{-1}$)**
H. pylori Trx1	25.14 ± 5.60	7.50 ± 0.83	17.86 ± 3.75	7.1×10^5
H. pylori Trx2	26.02 ± 3.96	0.78 ± 0.10	1.86 ± 0.41	7.0×10^4

Table B1: Kinetic constants for *H. pylori* Trx system using both *H. pylori* trx1 and *H. pylori* trx2 as substrates for HP TtxR.

B2: Materials and Methods

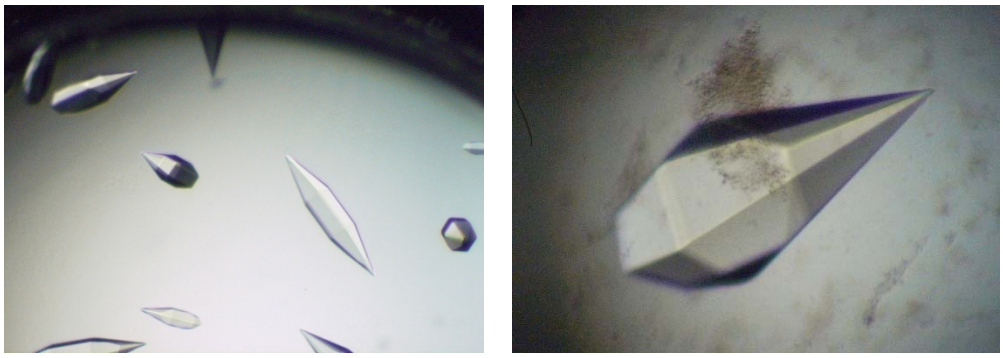
HP trx1 and HP trx2 were purified as described in sections 2.2.1 and 2.2.2 respectively. The purified proteins were then crystallized by vapor diffusion and microbatch methods.

B2.1: Crystallization of HP Trx1

Initial broad screening was done by vapor diffusion using commercial broad screens (Qiagen) at 20ºC but did not yield any crystals. Further broad screening by microbatch at 4ºC, yielded small needle-like crystals under the following conditions: 15% PEG 3350 and 0.2M sodium acetate buffer pH 4.6. Grid screening and use of additives did not improve the quality of the crystals.

B2.2: Crystallization of HP Trx2

Initial broad screening for *H. pylori* Trx2 was also done by vapor diffusion using commercial broad screens (Qiagen) at 20ºC. One condition (3.0M ammonium sulfate, 10% glycerol) from the ammonium sulfate broad screen yielded crystals. Grid screening and additive screening were performed based on these results to find the optimal growth conditions. The best condition was found to be 200 mM sodium acetate buffer (pH 5.0), 3.0M ammonium sulfate, and 10% glycerol (Figure B1). Spermine (200mM) and glycine (200mM) seemed to be the best additives, i.e. they improved the shape of the crystals.



A) **B)**

Figure B1: HP Trx2 crystals (before (left) and after (right) optimization). Crystals before optimization were small (0.1mm x 0.1mm x 0.2mm). After optimization the size increased to (0.3mm x 0.3mm x 0.4mm).

B3: Results and Discussion

The initial crystal screening for x-ray diffraction of crystals was done on beamline 08-1D.1 using Mar225CCD detector at the Canadian Light Source. HP Trx2 crystals from the optimized conditions (from grid screening and additive screening) showed diffraction spots to 4 Å (Table B2). Attempts were made to improve diffraction by dehydration and temperature annealing. For dehydration, 4µL of a super saturated solution of ammonium sulfate with 10% glycerol were added to Trx2 crystals and equilibrated against 100µL of the same dehydrating solution (sitting drop) overnight at 4°C before diffraction. Two temperature annealing methods were attempted on these crystals. The first attempt was made by blocking the cryo-stream for a few seconds and then cooling the crystal in the cryo-stream stream again. Another attempt was made by removing the crystal from the cryo-stream and placing it on the cryo-condition before re-looping it. None of these attempts improved diffraction.

Temperature (K)	100K
Wavelength (Å)	1.4
Space Group	P3
Unit Cell parameters (Å)	a= 80, b= 80, c= 571 α=β=90, γ=120
Resolution limits (Å)	50-4

Table B2: Data collection statistics for HP Trx 2

HP Trx1 crystals poorly diffracted (weak powder-like diffraction). Temperature annealing was used to try and improve diffraction, but diffraction did not improve.

REFERENCES

Adams P, Afonine P, Bunkóczi G, Chen V, Davis I, Echols N, Headd J, Hung L, Kapral G, Grosse-Kunstleve R, McCoy A, Moriarty N, Oeffner R, Read R, Richardson D, Richardson J, Terwilliger T, Zwart P (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213-221

Adler V, Yin Z, Tew K, Ronai Z (1999) Role of redox potential and reactive oxygen species in stress signaling. *Oncogene* **45:** 6104-6111

Ajay, Murcko M (1995) Computational methods to predict binding free energy in ligand-receptor complexes. *J Med Chem* **38:** 4953-4967

Akif M, Suhre K, Verma C, Mande S (2005) Conformational flexibility of Mycobacterium tuberculosis thioredoxin reductase: crystal structure and normal-mode analysis. *Acta Crystallogr D Biol Crystallogr* **61:** 1603-1611

Arner E, Holmgren A (2000) Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* **267:** 6102-6109

Arner E, Zhong L, Holmgren A (1999) Preparation and assay of mammalian thioredoxin and thioredoxin reductase. *Methods enzymol.* **300:** 226-239

Ashworth J, Havranek J, Duarte C, Sussman D, Monnat R, Stoddard B, Baker D (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441:** 656-659

Bahadur R, Zacharias M (2008) The interface of protein-protein complexes: analysis of contacts and prediction of interactions. *Cell Mol Life Sci* **65:** 1059-1072

Baker L, Raudonikiene A, Hoffman P, Poole L (2001a) Essential thioredoxin-dependent peroxiredoxin system from Helicobacter pylori: Genetic and kinetic characterization. *J Bacteriol.* **183:** 1961-1973

Baker N, Sept D, Joseph S, Holst M, McCammon J (2001b) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98:** 10037-10041

Bogan A, Thorn K (1998) Anatomy of hot spots in protein interfaces. *J Mol Biol* **280:** 1-9

Bradford M (1976) A rapid and senstive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding *Anal Biochem* **72:** 248-254

Bragg W (1913) *Proc Camb Phil Soc* 17: 43

Brim H, Osborne J, Kostandarithes H, Fredrickson J, Wackett L, Daly M (2006) Deinococcus radiodurans engineered for complete toluene degradation facilitates Cr(VI) reduction. *Microbiology* **152:** 2469-2477

Brunger A (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2:** 2728-2733

Buchanan B, Adamidi C, Lozano R, Yee B, Momma M, Kobrehel K, Ermel R (1997) Thioredoxin-linked mitigation of allergic responses to wheat. *Proc Natl Acad Sci U S A* **94:** 5372-5377

Bundle D, Sigurskjold B (1994) Determination of accurate thermodynamics of binding by titration microcalorimetry. *Methods Enzymol* **247:** 288-305

Bushweller J, Billeter M, Holmgren A, Wüthrich K (1994) The nuclear magnetic resonance solution structure of the mixed disulfide between Escherichia coli glutaredoxin(C14S) and glutathione. *J Mol Biol* **235:** 1585-1597

Carugo O, Argos P (1997) Protein-protein crystal-packing contacts *Protein Sci* 6**:** 2261-2263

Che Y, Brooks B, Marshall G (2006) Development of small molecules designed to modulate protein-protein interactions. *J Comput Aided Mol Des* **20:** 109-130

Chen R, Li L, Weng Z (2003) ZDOCK: An initial-stage protein-docking algorithm. *Poteins* **52:** 80-87

Chen R, Weng Z (2002) Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Poteins* **47:** 281-294

Clackson T, Wells J (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* **267:** 383-386

Cooper A (1999) Thermodynamic analysis of biomolecular interactions. *Curr Opin Chem Biol* **3:** 557-563

Correa P (1995) Helicobacter pylori and gastric carcinogenesis. *Am J Surg Pathol* **19 Suppl 1:** S37-43

Cox M, Battista J (2005) Deinococcus radiodurans - The consummate survivor. *Nat rev Microbiol* **3:** 882-892

Crowther R, Blow D (1967) A method of positioning a known molecule in an unknown crystal structure. *Acta Cryst.* **23**: 544-548

Dai S, Saarinen M, Ramaswamy S, Meyer Y, Jacquot J, Eklund H (1996) Crystal structure of Arabidopsis thaliana NADPH dependent thioredoxin reductase at 2.5 angstrom resolution. *J Mol Biol* **264:** 1044-1057

Daly M (2009) A new perspective on radiation resistance based on Deinococcus radiodurans. *Nat rev Microbiol* **7:** 237-245

Daly M, Gaidamakova E, Matrosova V, Vasilenko A, Zhai M, Venkateswaran A, Hess M, Omelchenko M, Kostandarithes H, Makarova K, Wackett L, Fredrickson J, Ghosal D (2004) Accumulation of Mn(II) in, Deinococcus radiodurans facilitates gamma-radiation resistance. *Science* **306:** 1025-1028

Davies G, Banatvala N, Collins C, Sheaff M, Abdi Y, Clements L, Rampton D (1994) Relationship between infective load of Helicobacter pylori and reactive oxygen metabolite production in antral mucosa. *Scand J Gastroenterol* **29:** 419-424

DeLano W, Ultsch M, de Vos A, Wells J (2000) Convergent solutions to binding at a protein-protein interface. *Science* **287:** 1279-1283

DeLano, W **(**2002a) The PyMol molecular graphics system. *DeLano Scientific,* San Carlos, CA.

DeLano W (2002b) Unraveling hot spots in binding interfaces: progress and challenges.*Curr Opin Struct Biol* **12:** 14-20

Dennis R, Micossi E, McCarthy J, Moe E, Gordon E, Kozielski-Stuhrmann S, Leonard G, McSweeney S (2006) Structure of the manganese superoxide dismutase from Deinococcus radiodurans in two crystal forms. *Acta Crystallogr F* **62:** 325-329

Devos D, Russell R (2007) A more complete, complexed and structured interactome. *Curr Opin Struct Biol* **17:** 370-377

Dixon D, Lindner D, Branchaud B, Lipscomb W (1979) Conformations and electronic structures of oxidized and reduced isoalloxazine. *Biochemistry* **18:** 5770-5775

Drenth J (2007) Principles of protein X-ray crystallography. *Springer New York*: Third edition

Eaton K, Morgan D, Krakowka S (1992) Motility as a factor in the colonisation of gnotobiotic piglets by Helicobacter pylori. *J Med Microbiol* **37:** 123-127

Elcock A, Sept D, McCammon J (2001) Computer simulation of protein-protein interactions. *J Phys Chem B* **105:** 1504-1518

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Epp O, Ladenstein R, Wendel A (1983) The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur J Biochem* **133:** 51-69

Gautier M, Lullien-Pellerin V, De Lamotte-Guery F, Guirao A, Joudrier P (1998) Characterization of wheat thioredoxin h cDNA and production of an active Triticum aestivum protein in Escherichia coli. *Eur J Biochem* **252 :** 314-324

Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31:** 3320-3323

Gromer S, Urig S, Becker K (2004) The thioredoxin system - From science to clinic. *Med Res Rev* **24:** 40-89

Guerois R, Nielsen J, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J Mol Biol* **320:** 369-387

Gustafsson T, Sandalova T, Lu J, Holmgren A, Schneider G (2007) High-resolution structures of oxidized and reduced thioredoxin reductase from Helicobacter pylori. *Acta Crystallogr D Biol Crystallogr* **63:** 833-843

Hassler M, Maley J (2004) A guide for isothermal titration calorimetry. *Saskatchewan Structural Sciences Centre*

He M, Smith A, Oslob J, Flanagan W, Braisted A, Whitty A, Cancilla M, Wang J, Lugovskoy A, Yoburn J, Fung A, Farrington G, Eldredge J, Day E, Cruz L, Cachero T, Miller S, Friedman J, Choong I, Cunningham B (2005) Small-molecule inhibition of TNF-alpha. *Science* **310:** 1022-1025

Holmgren A (1979) Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J Biol Chem* **254:** 9627-9632

Holmgren A, Morgan F (1976) Enzyme reduction of disulfide bonds by thioredoxin. The reactivity of disulfide bonds in human choriogonadotropin and its subunits. *Eur J Biochem* **70:** 377-383

Horn S, Lalowski M, Goehler H, Droege A, Wanker E, Stelzl U (2006) Huntingtin interacts with the receptor sorting family protein GASP2 - Review. *J Neural Transm* **113:** 1081-1090

Huo S, Massova I, Kollman P (2002) Computational alanine scanning of the 1 : 1 human growth hormone-receptor complex. *J Comput Chem* **23:** 15-27

Jackson R (1999) Comparison of protein-protein interactions in serine protease-inhibitor and antibody-antigen complexes: Implications for the protein docking problem. *Protein Sci* **8:** 603-613

Janin J (2007) The targets of CAPRI rounds 6-12. *Proteins* **69:** 699-703

Janin J, Chothia C (1990) The structure of protein-protein recognition sites. *J Biol Chem* **265:** 16027-16030

Jones S, Thornton J (1996) Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* **93:** 13-20

Joudrier P, Gautier M, de Lamotte F, Kobrehel K (2005) The thioredoxin h system: potential applications. *Biotechnol Adv* **23:** 81-85

Kabsch, W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* **26**: 795–800.

Karplus P, Schulz G (1989) Substrate binding and catalysis by glutathione reductase as derived from refined enzyme:substrate crystal structures at 2 Å resolution. *J Mol Biol* **210**: 163–180

Keegan R, Winn M (2008) MrBUMP: an automated pipeline for molecular replacement. *Acta Crystallogr D Biol Crystallogr* **64:** 119-124

Khoury G, Fazelinia H, Chin J, Pantazes R, Cirino P, Maranas C (2009) Computational design of Candida boidinii xylose reductase for altered cofactor specificity. *Potein Sci* **18:** 2125-2138

Kobayashi I, Tamura T, Sghaier H, Narumi I, Yamaguchi S, Umeda K, Inagaki K (2006) Characterization of monofunctional catalase KatA from radioresistant bacterium Deinococcus radiodurans. *J Biosci Bioeng* **101:** 315-321

Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci U S A* **99:** 14116-14121

Kortemme T, Baker D (2004a) Computational design of protein-protein interactions. *Curr Opin Chem Biol* **8:** 91-97

Kortemme T, Kin D, Baker D (2004b) Computational alanine scanning of protein-protein interfaces. *Sci STKE* **219***:* Pl2

Krnajski Z, Gilberger T, Walter R, Cowman A, Muller S (2002) Thioredoxin reductase is essential for the survival of Plasmodium falciparum erythrocytic stages. *J Biol Chem* **277:** 25970-25975

Kuriyan J, Krishna T, Wong L, Guenther B, Pahler A, Williams CJ, Model P (1991) Convergent evolution of similar function in two structurally divergent enzymes. *Nature* **352:** 172-174

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R, Thompson J, Gibson T, Higgins D (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23:** 2947-2948

Laskowski R, MacArthur M, Moss D, Thornton J (1993) PROCHECK - A Program to check the steriochemical quality of protein structures. *J Appl Cryst* **26:** 283-291

Lee, K, Lee, J (2008) Computational approaches to protein-protein docking. *Curr Proteomics* **5**: 10–19.

Lennon B, Williams C (1996) Enzyme-monitored turnover of Escherichia coli thioredoxin reductase: Insights for catalysis. *Biochemistry* **35:** 4704-4712

Lennon B, Williams C (1997) Reductive half-reaction of thioredoxin reductase from Escherichia coli. *Biochemistry* **36:** 9464-9477

Lennon B, Williams C, Ludwig M (1999) Crystal structure of reduced thioredoxin reductase from Escherichia coli: Structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. *Protein sci* **8:** 2366-2379

Lennon B, Williams C, Ludwig M (2000) Twists in catalysis: Alternating conformations of Escherichia coli thioredoxin reductase. *Science* **289:** 1190-1194

Leslie, A (1992) Joint CCP4 + ESF-EAMCB *Newsl Protein Crystallogr* **No. 26.**

Levin-Zaidman S, Englander J, Shimoni E, Sharma A, Minton K, Minsky A (2003) Ringlike structure of the Deinococcus radiodurans genome: A key to radioresistance? *Science* **299:** 254-256

Lin Y, Lu P, Tang C, Mei Q, Sandig G, Rodrigues A, Rushmore T, Shou M (2001) Substrate inhibition kinetics for cytochrome P450-catalyzed reactions.*Drug Metabol Despos* **29:** 368-374

Lo Conte L, Chothia C, Janin J (1999) The atomic structure of protein-protein recognition sites. *J Mol Biol* **285:** 2177-2198

Looger L, Dwyer M, Smith J, Hellinga H (2003) Computational design of receptor and sensor proteins with novel functions. *Nature* **423:** 185-190

Lundström J, Holmgren A (1993) Determination of the reduction-oxidation potential of the thioredoxin-like domains of protein disulfide-isomerase from the equilibrium with glutathione and thioredoxin. *Biochemistry* **32:** 6649-6655

Luthman M, Holmgren A (1982) Rat liver thioredoxin and thioredoxin reductase: purification and characterization. *Biochemistry* **21:** 6628-6633

MacPherson A (1998) Crystallization of biological macromolecules. *Cold Spring Harbor Laboratory Press*

Makarova K, Aravind L, Wolf Y, Tatusov R, Minton K, Koonin E, Daly M (2001) Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed from the perspective of comparative genomics. *Microbiol Mol Biol Rev* **65***:* 44-79

Marshall B, Barrett L, Prakash C, McCallum R, Guerrant R (1990) Urea protects Helicobacter (Campylobacter) pylori from the bactericidal effect of acid. *Gastroenterology* **99:** 697-702

Martin J, Bardwell J, Kuriyan J (1993) Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* **365:** 464-468

Massova I, Kollman P (1999) Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *J Am Chem Soc* **121:** 8133-8143

Mattevi A, Obmolova G, Sokatch J, Betzel, C, Hol W (1992) The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD^+ at 2.45 Å resolution. *Proteins* **13**: 336–351

Matthews B (1968) Solvent content of a crystal. *J Mol Biol* **33**: 491-497

Mattimore V, Battista J (1996) Radioresistance of Deinococcus radiodurans: Functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *J Bacteriol* **178:** 633-637

McCoy A, Grosse-Kunstleve R, Adams P, Winn M, Storoni L, Read R (2007) Phaser crystallographic software. *J Appl Cryst* **40**: 658-674

Miranda-Vizuete A, Damdimopoulos A, Gustafsson J, Spyrou G (1997) Cloning, expression, and characterization of a novel Escherichia coli thioredoxin. *J Biol Chem* **272:** 30841-30847

Moreira I, Fernandes P, Ramos M (2007) Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **68:** 803-812

Mulrooney S, Williams C (1997) Evidence for two conformational states of thioredoxin reductase from Escherichia coli: Use of intrinsic and extrinsic quenchers of flavin fluorescence as probes to observe domain rotation. *Protein Sci* **6:** 2188-2195

Murshudov G, Vagin A, Dodson E (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240-255

Mustacich D, Powis G (2000) Thioredoxin reductase. *Biochem J* **346 Pt 1:** 1-8

Navarro J, Gleason F, Cusanovich M, Fuchs J, Meyer T, Tollin G (1991) Kinetics of electron transfer from thioredoxin reductase to thioredoxin. *Biochemistry* **30:** 2192-2195

Nordberg J, Arner E (2001) Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. *Free Radic Biol Med* **31:** 1287-1312

Obiero J, Bonderoff S, Goertzen M, Sanders D (2006) Expression, purification, crystallization and preliminary X-ray crystallographic studies of Deinococcus radiodurans thioredoxin reductase. *Acta Crystallogr F* **62:** 757-760

Obiero J, Pittet V, Bonderoff S, Sanders D (2010) Thioredoxin System from Deinococcus radiodurans. *J Bacteriol* **192:** 494-501

Oblong J, Gasdaska P, Sherrill K, Powis G (1993) Purification of human thioredoxin reductase: properties and characterization by absorption and circular dichroism spectroscopy. *Biochemistry* **32:** 7271-7277

Oliveira M, Discola K, Alves S, Medrano F, Guimaraes B, Netto L (2010) Insights into the Specificity of Thioredoxin Reductase-Thioredoxin Interactions. A Structural and Functional Investigation of the Yeast Thioredoxin System. *Biochemistry* **49:** 3317-3326

Perrakis A, Sixma T, Wilson K, Lamzin V (1997) wARP: Improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr* **53:** 448-455

Pervin S (2006) Study of protein-protein interactions using human and Helicobacter pylori thioredoxin systems. *University of Sakatchewan thesis*

Pierce M, Raman C, Nall B (1999) Isothermal titration calorimetry of protein-protein interactions. *Methods* **19:** 213-221

Price N, Stevens L (1999) Fundamentals of Enzymology: The cell and molecular biology of catalytic proteins. *Oxford University Press*

Qin J, Clore G, Kennedy W, Huth J, Gronenborn A (1995) Solution structure of human thioredoxin in a mixed disulfide intermediate complex with its target peptide from the transcription factor NF kappa B. *Structure* **3:** 289-297

Rhodes G (2006) Crystallography made crystal clear. *Academic Press.* **Third edition**

Rudolph J (2007) Inhibiting transient protein-protein interactions: lessons from the Cdc25 protein tyrosine phosphatases. *Nature Rev Cancer* **7:** 202-211

Rupp B (2010) Biomolecular Crystallography: Principles, Practice, and Application to Strucural Biology. *Garland Science, Taylor and Francis Group, LLC*

Saboury A (2004) Binding isotherm determination by isothermal titration calorimetry - Interaction between Cu2+ and myelin basic protein. *J Therm Anal Calorim* **77:** 997-1004

Sambrook J, Russell D (2001) Molecular cloning: A laboratory manual. *Cold Spring Harbor Laboratory Press*

Sanders D (1998) Thioredoxin homodimers: Effects of mutation and oxidation on structure and dimer formation. *University of Arizona thesis*

Schulz G, Schirmer R, Sachsenheimer W, Pai E (1978) The structure of the flavoenzyme glutathione reductase. *Nature* **273:** 120-124

Seo H, Lee Y (2006) Occurrence of thioredoxin reductase in Deinococcus species, the UV resistant bacteria. *J Microbiol* **44:** 461-465

Sharp K (1998) Calculation of HyHel10-Lysozyme binding free energy changes: Effect of ten point mutations. *Proteins* **33:** 39-48

Sheinerman F, Norel R, Honig B (2000) Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol* **10:** 153-159

Suzuki H (2006) Protein-protein interactions in the mammalian brain. *J Physiol* **575:** 373-377

Thelander L (1967) Thioredoxin reductase. Characterization of a homogenous preparation from Escherichia coli B. *J Biol Chem* **242:** 852-859

Thorn K, Bogan A (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* **17:** 284-285

Tian B, Xu Z, Sun Z, Lin J, Hua Y (2007) Evaluation of the antioxidant effects of carotenoids from Deinococcus radiodurans through targeted mutagenesis, chemiluminescence, and DNA damage analyses. *Biochim Biophys Acta* **1770:** 902-911

Turnbull W, Daranas A (2003) On the value of c: Can low affinity systems be studied by isothermal titration calorimetry? *J Am Chem Soc* **125:** 14859-14866

Vagin A, Teplyyakov A (1997) MOLREP: an Automated Program for Molecular Replacement *J Appl Cryst* **30**: 1022-1025

Vakser I, Kundrotas P (2008) Predicting 3D structures of protein-protein complexes. *Curr Pharm Biotechnol* **9:** 57-66

Van den Berg P, Mulrooney S, Gobets B, Van Stokkum I, Van Hoek A, Williams C, Visser A (2001) Exploring the conformational equilibrium of E-coli thioredoxin reductase: Characterization of two catalytically important states by ultrafast flavin fluorescence spectroscopy. *Protein Sci* **10:** 2037-2049

Vassilev L, Vu B, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu E (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science 303***:** 844-848

Veine D, Mulrooney S, Wang P, Williams C (1998) Formation and properties of mixed disulfides between thioredoxin reductase from Escherichia coli and thioredoxin: Evidence that cysteine-138 functions to initiate dithiol-disulfide interchange and to accept the reducing equivalent from reduced flavin. *Protein Sci* **7:** 1441-1450

Waksman G, Krishna T, Williams CJ, Kuriyan J (1994) Crystal structure of Escherichia coli thioredoxin reductase refined at 2 A resolution. Implications for a large conformational change during catalysis. *J Mol Biol* **236:** 800-816

Wang P, Veine D, Ahn S, Williams C (1996) A stable mixed disulfide between thioredoxin reductase and its substrate, thioredoxin: Preparation and characterization. *Biochemistry* **35:** 4812-4819

Wang Y, Coulombe R, Cameron D, Thauvette L, Massariol M, Amon L, Fink D, Titolo S, Welchner E, Yoakim C, Archambault J, White P (2004) Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor. *J Biol Chem* **279:** 6976-6985

Weber P, Salemme F (2003) Applications of calorimetric methods to drug discovery and the study of protein interactions. *Curr Opin Struct Biol* **13:** 115-121

Weichsel A, Gasdaska J, Powis G, Montfort W (1996) Crystal structures of reduced, oxidized, and mutated human thioredoxins: Evidence for a regulatory homodimer. *Structure* **4:** 735-751

Wells J (1990) Additivity of mutational effects in proteins. *Biochemistry* **29:** 8509-8517

Wells J, McClendon C (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450:** 1001-1009

White O, Eisen J, Heidelberg J, Hickey E, Peterson J, Dodson R, Haft D, Gwinn M, Nelson W, Richardson D, Moffat K, Qin H, Jiang L, Pamphile W, Crosby M, Shen M, Vamathevan J, Lam P, McDonald L, Utterback T, Zalewski C, Makarova K, Aravind L, Daly M, Minton K, Fleischmann R, Ketchum K, Nelson K, Salzberg S, Smith H, Venter J, Fraser C (1999) Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. *Science* **286:** 1571-1577

Windle H, Fox A, Ni Eidhin D, Kelleher D (2000) The thioredoxin system of Helicobacter pylori. *J Biol Chem* **275:** 5081-5089

Wiseman T, Williston S, Brandts J, Lin L (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179:** 131-137

Ye J, Cho S, Fuselier J, Li W, Beckwith J, Rapoport T (2007) Crystal structure of an unusual thioredoxin protein with a zinc finger domain. *J Biol Chem* **282:** 34945-34951

Young L, Jernigan R, Covell D (1994) A role for surface hydrophobicity in protein-protein recognition. *Protein Sci* **3:** 717-729

Zhang Z, Bao R, Zhang Y, Yu J, Zhou C, Chen Y (2009) Crystal structure of Saccharomyces cerevisiae cytoplasmic thioredoxin reductase Trr1 reveals the structural basis for species-specific recognition of thioredoxin. *Biochim Biophys Acta* **1794:** 124-128

Zhang Z, Dorrell N, Wren B, Farthing M (2002) Helicobacter pylori adherence to gastric epithelial cells: a role for non-adhesin virulence genes. *J Med Microbiol* **51:** 495-502

Zhon L, Arner E, Ljung J, Aslund F, Holmgren A (1998) Rat and calf thioredoxin reductase are homologous to glutathione reductase with a carboxyl-terminal elongation containing a conserved catalytically active penultimate selenocysteine residue. *J Biol Chem* **273:** 8581-8591

Zhong L, Arner E, Holmgren A (2000) Structure and mechanism of mammalian thioredoxin reductase: The active site is a redox-active selenolthiol/selenenylsulfide formed from the conserved cysteine-selenocysteine sequence. *Proc Natl Acad Sci U S A* **97:** 5854-5859